Ogiuga musacocia O55633 synechocyst O95633 scenorhabdi O9xa09 streptomyce P74682 synechocyst O9cj27 lactococcus O81554 hepatitis c O33063 mycobacteri O83394 mouse mamma

Q46837 escherichia O18117 caenorhabdi Q91212 pseudomonas Q9rq19 escherichia Q9h049 homo sapien

Q99mp3 macaca fasc Q9du47 chiba virus Q30489 equus cabal O80770 arabidopsis

071037 human endog 071072 human endog 095283 homo sapien Q9hdc0 homo sapien

O9xx15 caenorhabdi O9xx16 caenorhabdi O9x18 mus musculu

29vt54 drosophila

Title: Perfect score:

Run on:

Sequence:

Scoring table:

Searched:

```
Hasulke S., Jinno Y.; "Isolation of an endogenous retrovirus gene, a "Isolation and localization of an endogenous retrovirus gene, a candidate gene for type 1 diabetes, and identification of a CA repeat marker at its locus.";
Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF064864, AAC689311; -.
EMBL; AF1048464, AAC689311; -.
                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-98449692; PubMed-9778243;
Lan M.S., Mason A., Coutant R., Chen.Q.Y., Vargas A., Rao J.,
Gomez R., Chalew S., Garry R., Maclaren N.K.;
"HERV-KIOS and immune-mediated (type 1) diabetes.";
Cell 95:14-16(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         153 AA; 17305 MW; 2B36E1656DF0C01A CRC64;
                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100.0%; Score 840; DB 4;
100.0%; Pred. No. 1.4e-81;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                    Created)
                                 Q9GMP3
Q9DU47
Q30489
Q80770
Q9VT54
Q9XX15
Q9XX16
                                                                                                      097003
055633
09N533
09XA09
P74682
09CJ27
081554
033063
                071037
                                                                                                                                                                                       046837
                                                                                                                                                                                                                                                                                                   PRT;
                112
66
66
67
77
110
112
122
123
123
                                                                                                                                                                                                                                                                                                                   10,
10,
12,
                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 100.
Matches 153; Conservative
01-MAY-1999 (TrEMBLrel.
01-NOV-1999 (TrEMBLrel.
ENVELOPE PROTEIN RIC-1.
                                                                                                                                                                                                                                                                                                                   01-MAY-1999 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
protein.
                491
472
117.5
87
82
81.5
79
76
76
76
76
75.5
                                                                                                                                 73.5
73
72.5
72.5
72
72
                                                                                                                                                                                               71.5
71
71
70.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Envelope |
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                 095280
                                                                                                                                                                                                                                                                                 RESULT
095280
095280 homo sapien
042043 unclassifie
099006 human endog
099010W3 homo sapien
0910W4 homo sapien
095284 homo sapien
095384 human endog
091044 homo sapien
095384 human endog
                                                                   (without alignments)
315.918 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                            84.0
1 MVTPVTWMDNPIEVYVNDSV......ECVANSVVILQNNEFGTIID 153
                                                            April 9, 2002, 16:56:51; Search time 70.84 Seconds
        GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                            473505 seqs, 146272329 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                             Listing first 45 summaries
                                           OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09UNW2
09UNW2
09HDBB
014273
095281
095281
095281
095281
095384
                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                        sp_unclassified:*
                                                                                                                                                                                                                                                                                                           sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                      sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                 sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                       sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                              US-09-490-700-36
                                                                                                                                                                                                                                                                                                                                                                sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                       sp_plant: *
                                                                                                                                                                                                                                                                                                sp_human:*
                                                                                                                                                                                                                                                               SPTREMBL 17:*
                                                                                                                                                                                                                                                                                          sp_fung1:*
                                                                                                                                                                                                                                                                                                                            sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0
100.0
96.5
995.8
93.3
93.1
92.7
92.7
92.7
92.7
92.5
92.6
92.5
92.5
```

Database :

ö

Gaps

Length 153;

61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120

> g ŏ

09yna7 human endog 095282 homo sapien 09yna5 human endog 09ukh7 homo sapien 09ukh3 homo sapien

Q9YNA7 095282 Q9YNA5 Q9UKH7 Q9UKH3

840 831 831 805 805 805 792 779 779 779 777 777 777 777

Score

Result No.

121

61

g ð 셤

0

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                  LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                    61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNCLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDHCPAKPEEEGMMINISIGKHYPPICLGRAPGC 60
                                                                 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                  1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human).
blankaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                  .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rao J.,
 Length 560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96.5%; Score 811; DB 4; Length 153; 96.7%; Pred. No. 1.7e-78;
                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A., Gomez R., Challew S., Garry R., Maclaren N.K.; "HERV-K10s and immune-mediated (type 1) diabetes."; cell 95:14-16(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153 AA; 17226 MW; 4632F483BFD9517A CRC64;
                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
Score 831; DB 12;
Pred. No. 5.5e-80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153 AA
                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                  121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
MEDLINE-98449692; Pubmed-9778243;
98.98;
99.38;
                                                                                                                                                                                                                                                                                                                                             01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AF084865; AAC68894.1;
Envelope protein.
SEQUENCE 153 AA: 17226 MW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                ENVELOPE PROTEIN RIC-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENVELOPE PROTEIN RIC-7
                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Sim
Matches 148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
 Query Match
                  Best Local
                                                                                                                                                                                                                                                                                                                                 09UNW3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09UNW2;
                                                                                                                                                                                                                                                                                                                Q9UNW3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9UNW2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S
                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                               RESULT
Q9UNW3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
Q9UNW2
                                                                                                                                                                                                                                                                                                                  qq
                                                                  à
                                                                                                                                                               qq
                                                                                                                                                                                               δy
                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                               g
                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                      MEDLINE-97386455: Pubbled-9244304;
Conrada., Weissmahr R.N., Boni J., Arcari R., Schupbach J., Mach B. | Nukuhrak
"A Thuman endogenous retroviral superantigen as candidate autolmmune
gene in type I diabetes.";
# -26-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and expression
61 LMPAVQNWLVEVPTVSPNSRFTYHWVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDILINE-99445825; Pubmed-10516026;
Toenjes R.R., Czauderna F., Kurth R.;
Genome wide screening, cioning, chromosomal assignment and es of full-length human endogenous retrovirus type K (HERV-K).";
J. Virol. 73:9187-9195(1999).
SEDBL: Y18890; CAB56604.1;
SECUENCE 560 AA; 63671 MW; 8A4565663901BC3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              153 AA; 17305 MW; 2B36E1656DF0C01A CRC64;
                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Score 840; DB 12;
100.0%; Pred. No. 1.4e-81;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             560 AA
                                                                 121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               153
                                               121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID
                                                                                                                                                                                                                                                                                                 Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF012337; AAC58457.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel, 13, 01-MAY-2000 (TrEMBLrel, 13, 01-MAR-2001 (TrEMBLrel, 16, ENV PROTEIN,
                                                                                                                                                                              042043;
01-JAN-1998 (TrEMBLrel. 05,
01-JAN-1998 (TrEMBLrel. 05,
01-NOV-1998 (TrEMBLrel. 08,
ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                                               unclassified Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 100.
Best Local Similarity 100.
Matches 153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                              NCBI_TaxID=35276;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Envelope protein.
SEQUENCE 153 AA
```

900060

RESULTION OF SECOND OF SEC

121

g

ο

õ g ο S

m

```
SIMILARITY: BÉLONGS TO 2N-FINGER CCHC TYPE FAMILY.
SIMILARITY: THE PROTEASE BELONGS TO THE PEPTIDASE FAMILY A1; ALSO
KNOWN AS THE EUKARYOTIC ASPARTYL PROTEASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Many human endogenous retrovirus K (HERV-K) proviruses are unique to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mucleotide sequence of human endogenous retrovirus genome related the mouse manager years of human endogenous retrovirus genome related the mouse manamary tumor virus genome.";

-1. Virol. 60:589-598(1986).

-1. SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE TRANSCENTPASS).

INTANSCRIPTASS).

INTERPO: IPR001037; Integrase_C.

InterPo: IPR003308; Integrase_C.

InterPo: IPR001584; RwaseH.

InterPo: IPR001584; RwaseH.

InterPo: IPR001584; RwaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                   Euteleostom1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECUENCE FROM N.A.
MEDLINE-94400989; PubMed=10469582;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1361;
                                                                                Craniata; Vertebrata; Eutele
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 792; DB 4; Length 1...
Pred. No. 2.1e-75;
Triminations 7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CEB91B3F407B9498 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 2294 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                TISSUE=LIVER;
MEDLINE=87036922; PubMed=3021993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RNA-directed DNA polymerase.
SEQUENCE 1361 AA; 153797 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00552; integrase; 1.
Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
                                                                             Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17, GAG-PRO-POL-ENV PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity 94.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                humans.";
Curr. Biol. 9:861-868(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-2000 (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PF00078; rvt; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF00665; rve; 1
                                                       Homo sapiens (Human)
                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSCRIPTASE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=9606;
                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Simi
Matches 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               о90кн9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
Q9UKH9
  DD BRANK RAY BRANK BRANK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 LMPAVQNWLVEVPTVSPNSRFTXHNVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jinno Y., Sugimoto J.;
"Human endogenous retrovirus HERV-K(II) and flanking sequences.";
Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AB047240; BABI1760.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö,
                                                                                                                                                                                                                                                                                                                                                   ..
0
                                                    Rao J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity 96.7%; Pred. No. 1.2e-77; Similarity 96.7%; Pred. No. 1.2e-77; Similarity 96.7%; Pred. No. 1.2e-77; Similarity 0; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                           Length 153;
                                                                                                                                                                                                                                                                                                                                             Indels
                MEDLINE-98449692; PubMed-9778243;
Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A.,
Gomez R., Chalew S., Garry R., Maclaren N.K.;
"HERV-KIDs and immune-mediated (type 1) diabetes.";
Cell 95:14-16(1998)
EMBL; AF084870; AAC68899.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27904 MW; 198F26D65ED56DDB CRC64;
                                                                                                                                                                                                          153 AA; 17406 MW; 90B91BCE0040C5A3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q14273 PRELIMINARY; PRT; 1361 AA. Q14273; 01-NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
                                                                                                                                                                                                                                                                                     Query Match 95.8%; Score 805; DB 4; L
Best Local Similarity 96.7%; Pred. No. 7.2e-78;
Matches 148; Conservative 0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          245 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 PKESKNTEVLVWEECVANSAVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 16, C
(TrEMBLrel. 16, I
(TrEMBLrel. 16, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-2001 (TrEMBLrel.
01-MAR-2001 (TrEMBLrel.
01-MAR-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              245 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                 Envelope protein.
SEQUENCE 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Simil
Matches 148;
```

Q9HDB8; 99нрв8 9

09HDB8

121 121

õ a

g

ð g ò SEQUENCE

ò 셤 ö g ò 7

RESULT

014273 ID 011 AC 011 DT 011

121

셤

ö

Gaps

ö

ö

0; Gaps

Length 153;

```
LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                           61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 LMPAVQNWLVEVPIVSPISRFTYHMVSGMSLRPRVNYLQDFPYQRSLKFRPKGKPCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MVIPVTWMDNPIEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                     1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 93.1%; Score 782; DB 4; Length 153; Best Local Similarity 94.1%; Pred. No. 2e-75; Matches 144; Conservative 0; Mismatches 9; Indels
                                                                        Score 784; DB 4; Length 15
Pred. No. 1.2e-75;
2; Mismatches 7; Indels
                    17391 MW; 4088880E06FB7677 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
MEDILNE-99449662; PubMed-9778243;
Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A., Gomez R., Chalew S., Garry R., Maclaren N.K.;
HERY-K10s and immune-mediated (type 1) diabetes.";
Cell 95:14-16(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             153 AA; 17360 MW; 28C01AD348C6D806 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
ENVELOPE PROTEIN RIC-3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 694 AA
                                                                                                                                                                                                                                                                                                PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                 121 PKESKNTEVLVWEECVDNSAVILQNSEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          153 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AF084866; AAC68895.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Human endogenous retrovirus K.
                                                                        93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, ENV PROTEIN.
                                                                        Query Match 93.3
Best Local Similarity 94.1
Matches 144; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens (Human)
   Envelope protein.
SEQUENCE 153 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Envelope protein.
SEQUENCE 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                             095281;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9YNA9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09YNA9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
Q9YNA9
                                                                                                                                                                                                                                                                                                                                                                                                                            095281
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                             095281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DT DT DT OS
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
   SS
                                                                                                                                                  ò
                                                                                                                                                                                      g
                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1855 LMPAVQNWLVEVPTVSPISRFTYHWVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKKI 1914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ENV.
Homo sapiens (Human).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ٦,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 93.9%; Score 789; DB 4; Length 2294; Best Local Similarity 94.8%; Pred. No. 7.9e-75; Matches 145; Conservative 1; Mismatches 7; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00443; G-patch; 1.
SMART; SM00343; ZnF_C2HC; 2.
SMOSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT_RETROV; 1.
ASPARTY1 protease; Hydrolase; RNA-directed DNA polymerase; Zinc-finger.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2294 AA; 257099 MW; 425C81FDDF8A80D0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDINE-98449622; PubMed-9778243;
MEDINE-98449622; PubMed-9778243;
GOMEZ R., Chalew S., Garry R., Maclaren N.K.;
"HERY-KIOS and immune-mediated (type 1) diabetes.";
Cell 95:41-16(1999).
EMBL; AF084869; AAC68898.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                153 AA
                                                     Asp_prot_retrov.duTPase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                    Interpro; IPR001428; dUTPase.
Interpro; IPR003322; Gag_pl0.
Interpro; IPR0000721; Gag_pz4.
Interpro; IPR000467; G_patch.
Interpro; IPR001037; Integrase_C.
Interpro; IPR03308; Integrase_zn.
                                    InterPro; IPR001969; Asp_protease
                                                                                                                                                                                                                                                     Pfam; PF00692; dUTPase; 1.
Pfam; PF01585; G-patch; 1.
Pfam; PF00337; Gag_pl0; 1.
Pfam; PF00607; gag_p24; 1.
Pfam; PF00552; integrase; 1.
Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF02025; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
AF164611; AAD51793.1; -. P16088; 1DUT.
                                                                                                                                                                                  Interpro; IPR002156; RNaseH.
Interpro; IPR001584; Rve.
Interpro; IPR000477; RVTse.
Interpro; IPR001878; Znf_CCHC.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00098; zf-CCHC; 2.
ProDom; PD000946; dUTPase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, ENVELOPE PROTEIN RIC-6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                        PF00665; rve; 1
PF00077; rvp; 1
                                                                                                                                                                                                                                                                                                                                                                                                                            PF00078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                        InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               095284
```

οy q ογ

Pfam; Pfam; Pfam; Pfam; Pfam;

Pfam;

Rao J.,

ö

Gaps

; 0

```
Length 699
 Score 779; DB 4; Length 69
Pred. No. 2.3e-74;
2; Mismatches 9; Indels
                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           514 AA
                                                                                                                                                          121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                        PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               669
                                                                                                                                                                                                                                                                                                                                         Human endogenous retrovirus.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=11827;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
 92.78;
92.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 92.7%;
Best Local Similarity 92.8%;
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ... wov-1996 (TremBLrel. 01, 01-NOV-1998 (TremBLrel. 08, POL/ENV GENE (FRAGMENT).
                                                                                                                                                                                                                                                                           01,
01,
08,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (TrEMBLrel. 01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                          (TrEMBLrel. 0 (TrEMBLrel. 0 (TrEMBLrel. 0
                           142; Conservative
                                                                                                                                                                                                                                                   PRELIMINARY;
                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                        01-NOV-1996 (
01-NOV-1996 (
01-NOV-1998 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1996
01-NOV-1996
  Query Match
Best Local S
Matches 142
                                                                                                                                                                                                                                                                                                                 ENV MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      069386;
                                                                                                                                                                                                                                                               069384;
                                                                                                                                                                                                                                                069384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        988690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        232
                                                                                                                                                                                                                         RESULT
Q69384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          069386
                                                                             q
                                                                                                                               g
                                                                                                                                                                                                                                                               ò
                                                                                                      ò
                                                                                                                                                          ò
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                               ö
                                                                             and
                                                                                                                                                                                                                                  LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                     1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "An almost-intact human endogenous retrovirus K on human chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Lenz J.;
"Many human endogenous retrovirus K (HERV-K) proviruses are unique
                                 SEQUENCE FROM N.A.
Toenjes R.R., Czauderna F., Kurth R.;
"Full-length human endogenous retrovirus type K (HERV-K) elements encoding Gag, Pol and Env proteins are localised on chromosomes 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Meese E.;
"Further characterization of the almost intact human endogenous
                                                                                                                                                                                               .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-99400999; Pubmed-10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                      Length 694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-99178257; Pubmed-10080172;
Mayer J., Sauter M., Racz A., Scherer D., Mueller-Lantzsch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
Reus K., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch
                                                                                                                                                                                              9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  retrovirus K on human chromosome 7.";
Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF164614; AAD517981;
EMBL; AF074086; AAR88168.1;
EMBL; AF074086; AAR88168.1;
EMBL; AF074086; AAD81098.1;
                                                                                                   (JUL-1998) to the EMBL/GenBank/DDBJ databases 132; CAA76880.1; -.
                                                                                                                  380.1; -.
78690 MW; 82F91825669CF25B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Envelope protein.
SEQUENCE 699 Aa; 79217 MW; 3844DA6D80D76DFC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (Human).
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; E
Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                   Score 779; DB 12;
Pred. No. 2.3e-74;
2; Mismatches 9;
                                                                                                                                                                                                                                                                                                                         121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                         Viruses; Retroid viruses; Retroviridae
                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                  92.7%;
92.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genet. 21:257-258(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAR-2001 (TrEMBLrel. 16, ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Curr. Biol. 9:861-868(1999).
                                                                                                                                                                               Best Local Similarity 92.8
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                         19..,
Submitted (JUL-1950,
EMBL; Y17832; CAA768f
            NCBI_TaxID-45617;
                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                 Q9UBU4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Meese E
                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                 167
                                                                                                                                                                                                                                                                                                                                                 227
g
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                  DDTTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDD
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                      ç
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                       61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                          1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-95074858; PubMed-7983704;
Loewer R., Toenjes R.R., Korbmacher C., Kurth R., Loewer J.;
Loewer R., Toenjes R.R., Korbmacher C., Kurth R., Loewer J.;
Lidentification of a Rev-related protein by analysis of spliced
transcripts of the human endogenous retroviruses HTDV/HERV-K.";
J. Virol. 69:141-149(1995).
EMBL: X82272; CAA57723.1; -.
SEQUENCE 699 AA; 79217 MW; 3844DA6D80D76DFC CRC64;
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 699;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 779; DB 12;
Pred. No. 2.3e-74;
2; Mismatches 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
SEQUELINE-97213939; Pubmed-9060628;
Tonjes R.R., Limbach C., Lower R., Kurth R.;
```

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Toenjes R.R., Czauderna F., Kurth R.;
Tenjis R.R., Czauderna F., Kurth R.;
"Fuli-length human endogenous retrovirus type K (HERV-K) elements
encoding Gag, Pol and Env proteins are localised on chromosomes 7 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                               1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                      .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 92.6%; Score 778; DB 12; Length 694; Best Local Similarity 92.8%; Pred. No. 3e-74; Matches 142; Conservative 2; Mismatches 9; Indels (
                                                                                                                                                                       92.6%; Score 778; DB 12; Length 514; 92.8%; Pred. No. 2.1e-74; 1ve 2; Mismatches 9; Indels
"Expression of human endogenous retrovirus type K envelope glycoprotein in insect and mammalian cells."; J. Virol. 71:2747-2756(1997).
                                                                    Toenjes R.R.;
Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases
EMBL; Y29887; CAA63481.1;
NON_TER 514 514
SEQUENCE 514 AA; 58200 MW; 86410F35B82DF231 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19.";
Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases
EMBL: Y17833; CAA76883.1; -.
SEQUENCE 694 AA; 78692 MW; 498A1D745366D0EE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                              694 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                             232 PKESKNTEVLVWEECVANSAVILQNNEFGTIID 264
                                                                                                                                                                                                                                                                                                                                            121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-2000 (TrEMBLrel. 13, ENV PROTEIN.
                                                                                                                                                                     Query Match 92.6
Best Local Similarity 92.8
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              09YNA7;
                                                                                                                                                                                                                                                                                                                                                                                                                                              09YNA7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               227
                                            2
                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
Q9YNA7
 ò
                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

Search completed: April 9, 2002, 17:06:31 Job time: 580 sec us-09-490-700-36.rpr

```
(without alignments)
295.730 Million cell updates/sec
                                                                                                                                         April 9, 2002, 16:54:06; Search time 39.41 Seconds
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                             OM protein - protein search, using sw model
                                                                                                                                                                                                                                            US-09-490-700-36
                                                                                                                                               Run on:
```

840 1 MVTPVTWMDNPIEVYVNDSV......ECVANSVVILQNNEFGTIID 153 219241 Total number of hits satisfying chosen parameters: 219241 segs, 76174552 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Perfect score: Scoring table: Sequence: Searched:

length: 0 length: 2000000000 sed Minimum DB Maximum DB

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

pir1:* pir2:* pir3:* pir4:* PIR_68:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

hypothetical prote hypothetical prote prostaglandin G/H alph-amylase (EC hypothetical prote env polyprotein probable transmemb probable polygalac env polyprotein -malate synthase (E probable transcrip hypothetical prote hypothetical prote cytokine receptor Pm5 protein - huma hypothetical prote molybdate-binding hypothetical prote glycinin G3 - soyb glycinin G3 precur spheroidin precurs retrovirus-related MHC class II histo hypothetical prote transcription fact conserved hypothet protein - fis probable membrane cyclooxygenase 1 Description SUMMARIES A53273 T02309 T126502 T126501 A35782 S76295 S39782 A36674 T36724 S76888 C86647 S69198 A44765 T45350 VCMVMM PYVZAM S21977 T25168 B83413 F83201 B55164 T51150 B86155 S26388 SYHOMA C65083 S04605 S11003 DB Query Match Length 365 602 1475 341 688 1003 1190 1262 251 295 387 Score 768.5 82 81.5 76 76 76 70.5 75.5 73.5 Result Š

unknown protein F2	hypothetical prote	amylopullulanase p	hypothetical prote	hypothetical prote	cytoplasmic trehal	probable alpha, alp	regulatory protein	hypothetical prote	env polyprotein -	polypeptide - hepa	DNA-directed RNA p	probable methyltra	hypothetical prote	forkhead transcrip	hypothetical prote
99296н	T16251	828669	T02386	T00273	H86024	S47739	RGBSCA	T29659	VCMVM	PC2219	S35548	B71510	T29490	S70357	T01922
7	7	~	~	N	7	7	7	7	٦	~	7	~	7	~	7
623	1274	1481	204	1520	549	549	214	355	688	876	1210	170	285	473	479
		<u> </u>	2	~	7	7	_	_	_			_	_	_	_
8.3	ж Э	 &	80	80	80	æ	80	80	80	8.1	8.1	8.0	8.	8	80.
69.5 8.3	69.5 8.3		o	8	68.5	68.5 8.		8 89			68 8.1			67.5 8.0	

ALIGNMENTS

RESULT

VCHUER VCHUER	Best Local Similarity 94.7%; Pred. No. 4.7e-69; Matches 142; Conservative 1; Mismatches 6; Indels 1.	
--	---	--

64 AVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG 123 셤 ò g à

g

MIC class II histocompatibility antigen DR alpha chain - horse (fragment)
C; Species: Equus caballus (domestic horse)
C; Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 21-Jan-2000
C; Accession: A53273
R; Albright, D.; Bailey, E.; Woodward, J.G.
Immunogenetics 34, 136-138, 1991
A; Title: Nucleotide sequence of a cDNA clone of the horse (Equus caballus) DRA gene. A; Reference number: A53273; MUID:91331619
A; Reference number: A53273
A; Status: preliminary
A; Molecule type: mRNA
A; Mesidues: 1-226 <ALB>
A; Cross-references: GB:M60100; NID:9164236; PIDN:AAA30956.1; PID:9164237

~

C: Genetics:

53

q ò Q

```
A;Cross-references: EMBL:AL023828; PIDN:CAA19456.1; GSPDB:GN00020; CESP:Y17G7B.10b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIDN:CAA19455.1; GSPDB:GN00020; CESP:Y17G7B.10a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical protein Y17G7B.10a - Caenorhabditis elegans
C;Species: Ceenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26501
                           C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---RSPAREIEECMSLS------PPVKKRMISEA- 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                272 PATTRWTDNP-----PREPAREIEECMSLS-----PPVKKRMISEA- 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59 GCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQ-DFSYQRSLKFRPKGKTCP 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -VEIPEVEKETVELPVLALNEVFI-----GESLSSRVSYYEIGINDAOMLKOKSSGIT-- 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        59 GCLMPAVQNWLVEVPTVSPNSRFTYHWVSGMSLRPRVNYLQ-DFSYQRSLKFRPKGKTCP 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -VEIPEVEKETVELPVLALNEVFI-----GESLSSRVSYYEIGINDAQMLKQKSSGIT-- 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PVT--WMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI---CLGRAP 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 PVT--WMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI---CLGRAP 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 514;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Map position: 2
A:Introns: 15/2; 34/3; 82/1; 146/3; 176/2; 262/2; 320/2; 445/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Map position: 2
A;Introns: 12/3; 60/1; 124/3; 154/2; 240/2; 298/2; 423/1
Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 2;
8.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-514 <WIL>
                                                                                                                                                                                                                                A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-492 <WIL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 9.0%; Score 76; DB 2
Best Local Similarity 26.0%; Pred. No. 8;
Matches 40; Conservative 21; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         118 KEIPKGSKNTEVLVW-----EECVANSVVIL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           359 -- ICTGTGSTS---WNFNINKLTEQCVQDLMKIV 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          118 KEIPKGSKNTEVLVW------EECVANSVVIL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              381 --ICTGTGSTS---WNFNINKLTEQCVODLMKIV 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Library, June 1998
                                                                                                                                             Library, June 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 76;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21;
                                                                                                                                                                                                                                                                                                                                                           A; Experimental source: clone Y17G7B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A, Cross-references: EMBL:AL023828; 8 A, Experimental source: clone Y17G7B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9.0%;
hypothetical protein Y17G7B.10b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                    A) Reference number: 220225
A) Accession: T26502
A) Statue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R; Smye, R. submitted to the EMBL Data A; Reference number: 220225 A; Accession: T26501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene: CESP:Y17G7B.10a
                                                                                                                                                                                                                                                                                                                                                                                                                   A; Gene: CESP: Y17G7B.10b
                                                                                                              R; Smye, R. submitted to the EMBL
                                                                                      C; Accession: T26502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local S:
Matches 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                       C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Probable transcription elongation factor [imported] - Arabidopsis thaliana
N.Alternate names: hypothetical protein F13P17.5
C. Species: Arabidopsis thaliana (mouse-ear cress)
C. Species: D. J. Lin, X.; Kaul. Sp. Sp. J. Lin, X.; Extransion T02309
A. Secretation: Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.
A. Accession: T02309
A. Status: Lanslated from GB/EMBL/DDBJ
A. Molecule type: DNA
A. Residues: 1-990 cROU>
A. Residues: 1-990 cROU>
A. Residues: Sembl. Accounting thaliana chromosome II BAC F13P17 genomic sequence.
A. Residues: S. Rountselv, S.D.: Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Conin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L. Bustre 402, 761-768, 1999
A. Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A. Accession: G84753
A. Status: preliminary
A. Molecule type: DNA
A. Residues: 1-990 cROU>
A. Title: Sequences and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A. Residues: 1-990 cSTO>
A. Cross-references: GB-ABC02093; NID:g3337352; PIDN:AAC27397.1; GSPDB:GN00139
C. Genetics: F13P17
A. Status Prilliniary
A. Residues: F13P17
A. Residues: GB-ABC02093; NID:g3337352; PIDN:AAC27397.1; GSPDB:GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Gene: F13p17.5; At2g34210
A;Map position: 2
A;Introns: 97/1; 177/3; 226/3; 280/3; 308/3; 332/2; 451/3; 503/3; 536/2; 552/3; 603/3;
C;Keywords: transcription factor
                                                                                                                                                                                                                                      5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
                     A;Map position: 20
C;Superfamily: class II histocompatibility antigen; immunoglobulin homology
F;97-162/Domain: immunoglobulin homology <IMM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  60 CLMPAVQNWLVEVP---TVSPNSRFTYHMVSGMSLRPRVNYL------QDFSYQRSL 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PMTPSSTSYLPTTPGGQAMTPGTDLDVMSLDIVNIFQFTDYVSLFFLCGHHQDGSCVVAL 917
                                                                                                                                                                                                                                                                                                                                                     -----FHYLPFLPSTEDV 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 PVTWMDNPIEVYVNDSVW---VPGPTDDRCPAKPEEEGMMINISIGYHYPPICL-GRAPG 59
                                                                                                                                                                                                                                                                                           2 VTPVTWMDN--PIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 108 KFRPKGKT------CPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTI 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      39;
                                                                                                                                                                                                                                      28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 990
                                                                                                                                                                      Query Match 9.8%; Score 82; DB 2; Length 226; Best Local Similarity 27.7%; Pred. No. 0.81; Matches 28; Conservative 16; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                 158 YDCKVEHWGLDEPLLKHWEFEAPT--PLSETTENVVCGLGL 196
                                                                                                                                                                                                                                                                                                                                                                                                                --CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSL 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.7%; Score 81.5; DB 2;
21.1%; Pred. No. 5.1;
tive 34; Mismatches 65;
                                                                                                                                                                                                                                                                                                                                | ||||: | |: ||::||: | ||: |
113 VIDVIWLKNGKPVTMGVSETVFL--PRDDQLFRK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 21.19
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
```

10;

Gaps

52;

10;

ζ

918

g ò g ò \sim

```
C;Accession: S3978.
R;Feng, L.; Sun, W.; Xia, Y.; Tang, W.W.; Chanmugam, P.; Soyoola, E.; Wilson, C.B.; H
A;Teite: Cloning two isoforms of rat cyclooxygenase: differential regulation of their
A;Reference number: S39782; MUID:94099619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N.Alternate names: nuclear factor 3A, hepatic
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 16-Jul-1999
C;Accession: A36674; S15651
R;Lai, E.; Prezioso, V.R.; Smith, E.; Litvin, O.; Costa, R.H.; Darnell Jr., J.E.
Genes Dev. 4, 1427-1436, 1990
A;Title: HNF-3A, a hepatocyte-enriched transcription factor of novel structure is reg
A;Reference number: A36674; MUID:91032994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: mRNA
A; Residues: 1-466 <LAI>
A; Residues: 1-466 <LAI>
A; Cross-references: GB:X55955; NID:956369; PIDN:CAA39418.1; PID:956370
C; Superfamily: transcription factor HNF-3; fork head DNA-binding domain homology
C; Keywords: DNA binding; liver; nucleus; transcription factor
F;124-288/Region: DNA binding #status predicted
F;170-261/Domain: fork head DNA-binding domain homology <FHD>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                           C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 10-Dec-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79 WLRSSLRPSPSFTHFLLTHGYWIWEFVNATFIREVLMGWVLTVRSNLIPSPPTYNTAHDY 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       322 PGPAASPQTLDHSGATATGGGSELKSPASSSAPPISSGPGGWICTPLSPTWLAPHESQLH 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 PPPVLLTDA-GVPSPVIPCCYYPCQNQGVCVRFGLD-HYQCDCTRTGYSGPNCTIPEIWT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PID:9460556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11 PIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLG---RAPGCLMPAVQN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Molecule type: mRNA
A;Residues: 1-602 <FENS
A;Cross-references: GB:S67721; NID:g460555; PIDN:AAB29400.1; PID:g46055
C;Superfamily: human prostaglandin-endoperoxide synthase; EGF homology
F;38-71/Domain: EGF homology <EGF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----NSRFTYHMVSGMSLRPRVNYL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---VPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  382 LKGAPHYSFNHPFSINNL--MSSSEQQHKLDFKAYEQALQYSPYGATLPASLPLG 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23 PGP----TDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNWLVE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8.8%; Score 74; DB 2; Length 466; 25.2%; Pred. No. 12; Live 13; Mismatches 57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99 ---QDFS----YQRSLKFRPKGKTCPKEIPKGSKNTEVL 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 74.5; I
Pred. No. 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            transcription factor HNF-3A, hepatic - rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         68 WLVEVPTVSP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                 cyclooxygenase 1 - rat
                                                  1 | | :
311 ESPSGKE 317
119 EIPKGSK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Accession: S39782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
Op
    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                              C; Accession: A35782
R; Gorman, D.M.; Itoh. N.; Kitamura, T.; Schreurs, J.; Yonehara, S.; Yahara, I.; Arai, K. Proc. Natl. Acad. Sci. U. S.A. 87, 5499-5463, 1990
A; Title: Cloning and expression of a gene encoding an interleukin 3 receptor-like protein A37182
A; Accession: A35782
A; Molecule type: mRNA
A; Nesidues: 1-896 GOR>
A; Residues: 1-896 GOR>
A; Cross-references: GB:MJ4397; NID:g191821; PIDN:AA37204.1; PID:g309101
C; Comment: Mouse high-affinity IL-5, GM-CSF, and one class of high-affinity IL-3 receptor C; Superfamily: interleukin-3 receptor: beta chain; cytckine receptor; duplication; transmembrane protein
F; 1-22/Domain: signal sequence #status predicted <SIG>
F; 23-896/Product: cytckine receptor common beta chain #status predicted <MAT>
F; 23-895/Product: cytckine receptor homology <CRST>
F; 23-895/Product: cytckine receptor common beta chain #status predicted <CRST>
F; 23-895/Product: cytckine receptor common
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pypounciar process.

Specifical Symbolication of the process of th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Residues: 1-318 <KAN>
A;Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAA10147.1; PID:d101079
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                             Species: Mus musculus (house mouse)
Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOGRIDAFAGDGILLFG------EALVLDIPLGRSYRLYPPYPLDCQGY--GLI 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MPAVQ-NW--LVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPK 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 VPG-PTDDRCPAKPEEEGMMINISIG-YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNS 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIG--YH-YPPI---CLGRAPGCL 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       846 E-TEDLVQDLSVKKFPYQPMPQAPAIQFFKSLKHQDYLSLPPWDNSQSGKVC 896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80 RFTYHMVSGMSLR-----PRVNYLQDFSYQRSLKFRP----KGKTC 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9.0%; Score 75.5; DB 2; Length 318; 27.6%; Pred. No. 5.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9.0%; Score 76; DB 1; Length 896; 25.9%; Pred. No. 16; ive 18; Mismatches 39; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synechocystis sp. (strain PCC 6803)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;253-434/Domain: cytokine receptor homology <CRS2>
F;442-463/Domain: transmembrane *status predicted <TMM>
F;464-896/Domain: intracellular *status predicted <INT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51;
                                                                                                                        cytokine receptor common beta chain precursor - mouse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Reference number: S74322; MUID:97061201
A;Accession: S76295
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 9.0°
Best Local Similarity 25.9°
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

딤

```
A; Molecule type: mRNA
A; Residues: 1-602 <KIT>
A; Residues: 1-602 <KIT>
A; Cross-references: EMBL:U18060; NID:g603051; PIDN:AAA85823.1; PID:g603052
B; Kitzler, J.; Hill, E.; Hardman, R.; Reddy, N.; Philpot, R.; Eling, T.E.
Arch. Biochem. Blophys. 316, 856-863, 1995
A; Title: Analysis and quantitation of splicing variants of the TPA-inducible PGHS-1 of A; Reference number: S69199; MUID:95168876
7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                 A,Status: preliminary
A,Molecule type: DNA
A,Residues: 1-365 <STO>
A,Residues: GB:AE005176; NID:912723031; PIDN:AAK04277.1; GSPDB:GN00146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                prostaglandin G/H synthase 1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 24-aug-1996 #sequence_revision 13-Mar-1997 #text_change 10-Dec-1999
C;Accession: S69198; S69199
R;Kitzler, J.W.
submitted to the EMBL Data Library, December 1994
A;Reference number: S69198
A;Reference number: S69198
A;Accession: S69198
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           || :: || : ||: || || WLRS--SLRPSPSFTHFILLTHGYWIWEFVNATFIREVLMRLVITVRSNLIPSPPTYNTAH 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81 FTYH--MVSGMSLRPRV-NYLQD-----FSYQRSLKFRPKGKTCPKEIPKGSKNTEVL 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21 PPPVLLTDA-GVPSPVNPCCYYPCQNQGVCVRFGLD-HYQCDCTRTGYSGPNCTIPEIWT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31 PAKPEE---EGMMINISIGYHYPPICLGRAPGCLMPAVQN------WLVEVPTVSPNSR 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11 PIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLG---RAPGCLMPAVQN 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 2; Length 602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 365;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Cross-references: EMBL:U18060
A:Wote: only a part of the nucleic acid sequence is shown
C;Superfamily: human prostaglandin-endoperoxide synthase; EGF
C;Reywords: alternative splicing
F;38-71/Domain: EGF homology <EGF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Gene: ybiD
C;Superfamily: conserved hypothetical protein HI0365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
8.6%; Score 72.5; DE
Best Local Similarity 25.2%; Pred. No. 13;
Matches 36; Conservative 22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8.6%; Score 72.5; D. 21.7%; Pred. No. 23; tive 20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WLVEVPTVSPNSRFTYHM------VSGMSLR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             324 FYDVLKKNGINCVVRQEHGTDID 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        131 VWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Experimental source: strain IL1403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Molecule type: mRNA
A;Residues: 61-602 <KI2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       861698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
S76888
hypothetical protein s110446 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
G;Species: Synechocystis sp.
G;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76888
F;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Oymura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule fype: DNA
A; Residues: 1-1072 «KAN»
A; Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAA18800.1; PID:g165389
A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C; Superfamily: Synechocystis hypothetical protein s110446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hypothetical protein ybiD [imported] - Lactococcus lactis subsp. lactis (strain IL1403) C;Species: Lactococcus lactis subsp. lactis
                                                                                                                                                                        probable membrane protein - Streptomyces coelicolor C;Species: Streptomyces coelicolor C;Species: Streptomyces coelicolor C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999 C;Accession: T36724 R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. abumitted to the EMBL Data Library, June 1999 A;Reference number: 221612 A;Reference number: 221612 A;Accession: T36724 A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Cross-references: EMBL:AL079308; PIDN:CAB45222.1; GSPDB:GN00070; SCOEDB:SCH69.25c A; Experimental source: strain A3(2) C; Genetics: A; Genetics: Group B; Genetics: Group 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Э,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .,
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           130 MVPP----DGPWAELYLFELDGVAIGMTSPATIVCPTGYFPEHLSRRIAWIKNGFFTNPI 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53 CLGRAPG---CLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKF 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  186 GNGLAPNQKEVLLPWLENLKGNLMR-NPQNEDLAGRVSGMI----DDYIQDLGVTTSNVF 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MVTPVTWMDNP-IEVYVND----SVWVPGPTDDRCPAK--PEEEGMMIN-ISIGYHYPPI 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 DSVWVPG-----PTDDRC----PAKPEEEGMMINISIGYHYPPICLGRAPGCLMPA 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.8%; Score 73.5; DB 2; Length 297; ilarity 33.9%; Pred. No. 7.9; Conservative 6; Mismatches 18; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 2; Length 1072; 40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              71; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8.7%; Score 73; DB 28.2%; Pred. No. 40; tive 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RPKGKTCPKEIPKGSKNTEVLVWEECVAN 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 OPTTKNLPFGIPLAPPPLNALYPAEAVON 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Reference number: S74322; MUID:97061201
A;Accession: S76888
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 28.2
Matches 42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
Matches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C86647
```

g

à

į

ò g ò 요 δ

```
102
                   op
                                                                                                                                                                                                                                                          C;Accession: A44765
R;Melasniemi, H.; Paloheimo, M.; Hemioe, L.
A;Gen. Microbiol. 136, 447-454, 1990
A;Title: Nucleotide sequence of the alpha-amylase-pullulanase gene from Clostridium ther A;Reference number: A44765; MUID:90362027
                                                                                                                                                                                alpha-amylase (EC 3.2.1.1) / alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) pred
C;Species: Thermoanaerobacter thermohydrosulfuricus
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                  A.Status: preliminary
A.Molecule type: DNA
A.Molecule type: 1. PID:g144727
C.Function: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A.Posthway: glycogen/starch degradation
C.Superfamily: flbronectin type III repeat homology
C.Superfamily: flbronectin type III repeat homology
C.Keywords: glycosidase; hydrolase; polysaccharide degradation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               T43330

T43330

T43330

T43330

C; Species: Mycobacterium leprae

C; Accession: T45350

A; Reference number: 216918

A; Reference number: 216918

A; Molecule type: DNA

A; Residues: 1-341 cPaR>

A; Cross-references: EMBL:29494; PIDN:CAB16678.1

A; Experimental source: cosmid B57

C; Genetics:

A; Note: MLCB57.38

C; Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1080 NAVYNKVEGVEISKDWTWVDAQYVGDSGNNDKYMAKFVPDMVGTWEYIMRFSSNQGHDWT 1139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1031 THVIGVNNPVEVYA--EVWAQGLTD-----KPGQGENMI-AQLGYRYIGDTVGDA---VY 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63 PAVQN-----SLRPRVNYLQDFS 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 TPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLM 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 TPVTWMDNPIEVYVNDSVWVPGPTD--DRCP---AKPEEEGMMINI-----SIGYHYP 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8.6%; Score 72.5; DB 2; Length 1475; 28.6%; Pred. No. 65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 341;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
  DF-----SYQRSLKFRPKGKTCPKEIPKGSKNTEVL 130
                                 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match

8.6%; Score 72; DB 2;
Best Local Similarity 25.5%; Pred. No. 13;
Matches 24; Conservative 13; Mismatches 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 28.68;
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | : | |||
| 1140 YTKG----PDGKT 1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103 YQRSLKFRPKGKT 115
                                                                                                                                 14
100
ò
                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

Search completed: April 9, 2002, 16:58:10 Job time: 244 sec

:		
1		
17.		
F		
1		
	v .	
l.	•	
	T-	
P.		
		•
e V		

1		
h.		
4		
a a		
	型 (3+ (名) 1 (3+ (3+ (3+ (3+ (3+ (3+ (3+ (3+ (3+ (3+	
3 14		
i i		
a Se [‡]	t ve se	
	* * * * * * * * * * * * * * * * * * *	
i.		

```
GENERAL INFORMATION:
APPLICANT: MOYER, Richard W.
APPLICANT: MOYER, Richard L.
APPLICANT: Gruid1, Michael E.
TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
                                                          Sednence Sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTE READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FO COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435
                                           US-08-808-720-7
US-08-53-894-10
US-08-53-894-10
US-08-513-894-10
US-08-468-570-87
US-08-20-655A-87
US-08-868-373-14
US-08-470-139-26
US-08-470-299-10
US-08-54-894-12
US-08-08-64-888-100
US-08-08-64-888-100
US-08-08-91-90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION UNDRER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REGISTRATION NUMBER: 31,794
RECISTRATION NUMBER: UF114.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-775-8100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 6, Application US/07991867B Patent No. 5476781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 1003 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 904-372-5800 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      904-372-5800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-991-8678-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 1
US-07-991-867B-6
$25255
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
$2525
    8001284888888844444
800128489786012844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appli
Appl
Appl
Appl
Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
Appli
Appl
Appl
                                                                                                                                                                                                                                                                                                            April 9, 2002, 16:53:21; Search time 34.73 Seconds (without alignments) 99.136 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Appli
Appli
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 15, A Sequence 25, A Sequence 10, A Sequence 10, A Sequence 10, A Sequence 451, Sequence 2, Ap Sequence 12, A Sequence 14, A Sequence 15, A Sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 6, Al
Sequence 99, A
Sequence 7, Al
Sequence 15, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MVTPVTWMDNPIEVYVNDSV......ECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence 6,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1
Sequence 1
Sequence 3
Sequence 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                                                GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-544-332-6
US-08-687-076-99
US-08-220-151-15
US-08-413-118-15
US-09-078-173-25
US-09-078-173-25
US-09-537-357-32
US-09-537-357-32
US-09-537-357-32
US-08-313-288E-10
US-08-313-288E-10
US-08-313-288E-10
S-08-313-288E-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-258-754-451
US-08-218-686-2
US-08-466-242-2
US-07-862-0218-12
US-08-313-288B-12
PCT-US93-03164-12
US-08-688-988-40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-218-265-14
US-08-521-872-14
US-08-590-399-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-540-804-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-808-720-3
US-08-808-720-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     212252 segs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                    OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-490-700-36
840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  807
2409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  711.7
711.7
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
711.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64.5
64.5
64.5
64.5
64
64
64
64
64
63.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2525250
2525250
25253350
25253350
25253350
25253350
25253350
25253350
25253350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
2525350
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
252530
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
25250
                                                                                                                                                                                                                                                                                                                           Run on
```

```
COUNTRY:
                                           JS-08-544-332-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                           5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5
                                                                                     60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System
CORRESPONDENCE ADDRESS:
Length 1003;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8.5%; Score 71.5; DB 1; Length 1003; 25.0%; Pred. No. 11; ive 12; Mismatches 37; Indels 29
                                         37; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    441 CLKPKVPKNLRLWGWILDCDT----SRFIKHMADGSDDLDLDVRLN----
                                                                                                                            441 CLKPKVPKNLRLWGWILDCDT---SRFIKHMADGSDDLDLDVRLN-
                                                                                                                                                                                                         : | |: | : | : | 483 -RNDICLKQAIKQHYTNVIILEYANTYPNCTLSLGNNRFNNVFD 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

FILING DATE: 19-40G-1993

CLASSIFICATION: 435

PRIOR APPLICATION THER: 05 07/827,658

FILING DATE: 30-JAN-1992

PRIOR APPLICATION NUMBER: 05 07/657,584

FILING DATE: 19-FEB-1991

APPLICATION NUMBER: 05 07/657,584

FILING DATE: 19-FEB-1991

ATCHREY/AGENT INFORMATION:

NAMM: SAILWANCHIN, DAVIG R

REGISTRATION NUMBER: 31,794

REGISTRATION NUMBER: 31,794

REGISTRATION NUMBER: 31,794

TELEPHONE: (904) 375-8000

INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       483 -RNDICLKQAIKQHYTNVIILEYANTYPNCTLSLGNNRFNNVFD 525
                                                                                                                                                                             111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
; Score 71.5; D
; Pred. No. 11;
12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                   Sequence 6, Application US/08107755A Patent No. 5721352 GENERAL INFORMATION:
8.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 1003 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 25.09
Matches 26; Conservative
                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
Query Match
Best Local Similarity
Matches 26; Conserval
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  U.S.A.
                                                                                                                                                                                                                                                                                          RESULT 2
US-08-107-755A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-107-755A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                             õ
```

```
.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
Sequence 6, Application US/08544332
Patent No. 593577
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Hall, Richard W.
APPLICANT: Hall, Michael E.
APPLICANT: Gruidl, Michael E.
APPLICANT: Gruidl, Michael E.
APPLICANT: Gruidl, Michael E.
APPLICANT: Gruidl, Michael E.
CONRESPONDENCES: 77
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : | | : | : | : | : | - : | : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : | - : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           441 CLKPKVPKNLRLWGWILDCDT----SRFIKHMADGSDDLDLDVRLN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION: 435
CLASSIFICATION: 435
PRIOR APPLICATION UNDER: 45
PRIOR APPLICATION NUMBER: 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION NUMBER: US 07/827,685
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: US 07/857,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: COASTAN HANDER OF THE TRANSENT OF THE TRANSENT OF THE TRANSENT ON THE TRANS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Bencen, Gerard H. REGISTRATION NUMBER: 35,746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 25.0%;
Matches 26; Conservative 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1003 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 904-372-5800 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      904-372-5800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: protein US-08-544-332-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32606
```

us-09-490-700-36.rai

```
APPLICANT: Paolett1, Enzo
APPLICANT: Limbach, Keith J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --RPRVNYLQ---DFSYQRSLKFRPKGKTC----PKEIPK 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            376 HESPVSYKRFTKSVAPDVYYP-----PTVSVTFADTRAICDVKCVPRDGISLMWKI 426
                                                                                                                                                   378 GNYHLPKAMSADILITGPCIERPGLVNIQSMCDISETDG----PVSYTCQTIGYPPILP- 432
                                                                                                  83 --YHMVSGMSL-----PKEIPK 122
                                                 327 HESPVSYKRFTKSVAPDVYYP-----PTVSVTFADTRAICDVKCVPRDGISLMWKI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48 HYPPICLGR-----APGCLMPAVQNWLVEVPTVSPNSRFT-------82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 TPVTWM--DNPIE-VYVND--SVWVPG------PTDDRCPAKPEEEGMMINISIGY 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 533;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8.0%; Score 67.5; DB 1; Length 5
23.2%; Pred. No. 13;
tive 20; Mismatches 54; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
     48 HYPPICLGR-----APGCLMPAVQNWLVEVPTVSPNSRFT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-333
TELEFAX: (212) 840-0712
TELEFAX: 425066 CURTMS
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Curtis, Morris & Safford
                                                                                                                                                                                                                                 | :|:| : : : : : 433 GFYDTQVYDASPEIVSESMLVSVVAVI 459
                                                                                                                                                                                                    123 GSKNTEV-----LVWEECVANSVVIL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --
--
                                                                                                                                                                                                                                                                                                                                                                          Sequence 15, Application US/08220151 Patent No. 5529780 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US/00
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET: 530 Fifth Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH: 533 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 peptide
N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 23.2%
Matches 48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: NUCL
TITLE OF INVENTION: CANI
NUMBER OF EXQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUTLIS, MOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---YHMVSGMSL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE:
FRAGMENT TYPE:
                                                                                                                                                                                                                                                                                                                              RESULT 6
US-08-220-151-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             á
                                                 g
                                                                                                    δλ
                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: AUDONNET, Jean-Christophe F.
APPLICANT: BAUDU, Philippe G.
APPLICANT: RIVIERE, MICHEL A.
TITLE OF INVENTION: RECOMBINANT VACCINE CONTAINING FELINE HERPES VIRUS TYPE
TITLE OF INVENTION: 1, PARTICULARLY FOR TREATING FELINE INFECTIOUS
TITLE OF INVENTION: 0, PARTICULARLY FOR TREATING FELINE INFECTIOUS
TITLE OF INVENTION: 2, PARTICULARLY FOR TREATING FELINE INFECTIOUS
FILE REFERENCE: AUDONNET
CURRENT APPLICATION NUMBER: US/09/080,044
CURRENT APPLICATION NUMBER: PCT/FF96/01830
EARLIER APPLICATION NUMBER: 95/14450
EARLIER PILING DATE: 1996-11-30
SERLIER PILING DATE: 1995-11-30
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | :|| ||| |:| |||:| |||:| 272 TKLTWYLDGKPIERQYISDTASVWIDGLITRSSVLAIPTTETDSEKPD----IRCDLEW 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 TPVTWM--DNPIE-VYVND--SVWVPG------PTDDRCPAKPEEEGMMINISIGY 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50 PPICLGRAPGCL--MPA------VQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVN 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 4; Length 473;
                                                                   APPLICANT: Ruvkun, Gary
APPLICANT: Kimura, Koutarou
APPLICANT: Kimura, Koutarou
APPLICANT: Ogg, Scott
APPLICANT: Ogg, Scott
APPLICANT: Paradis, Suzanne
APPLICANT: Paradis, Suzanne
APPLICANT: Pissenbaum, Heidi
APPLICANT: Morris, Jason
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
FILE REFERENCE: ONOR6/A51001
CURRENT APPLICATION NUMBER: US/08/857,076C
CURRENT APPLICATION NUMBER: US/08/857,076C
NUMBER OF SEO ID NOS: 114
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
8.0%; Score 67.5; DB
Best Local Similarity 23.2%; Pred. No. 12;
Matches 48; Conservative 20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match

8.0%; Score 67.5; DB

Best Local Similarity 30.7%; Pred. No. 11;

Matches 27; Conservative 13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97 YLQDFSYQRSLKFRPKGKTCPKEIPKGS 124
Sequence 99, Application US/08857076C Patent No. 6225120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 7, Application US/09080044 Patent No. 6074649 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; ORGANISM: Feline herpesvirus 1
US-09-080-044-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Homo sapiens
US-08-857-076-99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

8 g ò 셤 14;

å

```
10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
US-09-078-173A-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-473-446-15
                                                          US-08-473-446-15
                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14;
                                                                                                                                                                                                                                       APPLICANT: PAOLETTI, ENZO
APPLICANT: LIMBACH, KEITH J.
TITLE OF INVENTION: NOCLECOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANTINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | :|| | || ::| || || ::| 321 TKLTWYLDGKPIERQYISDTASVWIDGLITRSSVLAIPTTETDSEKPD-----IRCDLEW 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    376 HESPVSYKRFTKSVAPDVYYP-----PTVSVTFADTRAICDVKCVPRDGISLMWKI 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83 --YHMVSGMSL-----PKEIPK 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    427 GNYHLPKAMSADILITGPCIERPGLVNIQSMCDISETDG----PVSYTCQTIGYPPILP- 481
427 GNYHLPKAMSADILITGPCIERPGLVNIQSMCDISETDG----PVSYTCQTIGYPPILP- 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 TPVTWM--DNPIE-VYVND--SVWVPG------PIDDRCPAKPEEEGMMINISIGY 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 533;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: US/08/413,118 FILING DATE: 29-MAR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              48 HYPPICLGR----APGCLMPAVQNWLVEVPTVSPNSRFT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8.0%; Score 67.5; D
23.2%; Pred. No. 13;
Live 20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,151
FILING DATE: 30-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/POCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (21,2),840-3333
                                                             | :|:| : | :: | :: 482 GFYDTQVYDASPEIVSESMLVSVVAVI 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       482 GFYDTQVYDASPEIVSESMLVSVVAVI 508
                                     123 GSKNTEV-----LVWEECVANSVVIL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            123 GSKNTEV-----LVWEECVANSVVIL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                               UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                              Sequence 15, Application US/08413118 Patent No. 5688920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 8.0%
Best Local Similarity 23.2%
Matches 48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-413-118-15
                                                                                                                                                           US-08-413-118-15
g
                                       δ
                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
14;
                                                                          APPLICANT: PAOLETII, ENZO
PEPLICANT: LIMBACH, KEITH J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                376 HESPVSYKRFTKSVAPDVYYP-----PTVSVTFADTRAICDVKCVPRDGISLMWKI 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --YHMVSGMSL-----PKEIPK 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                427 GNYHLPKAMSADILITGPCIERPGLVNIQSMCDISETDG----PVSYTCQTIGYPPILP- 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 TPVTWM--DNPIE-VYVND--SVWVPG------PTDDRCPAKPEEEGMMINISIGY 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 533;
                                                                                                                                                                                                                                                                                                                                                                                 COMPOUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.30
APPLICATION NUMBER: US/08/473,446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48 HYPPICLGR-----APGCLMPAVQNWLVEVPTVSPNSRFT-
                                                                                                                                                                                                                                       P.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: PROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8.0%; Score 67.5;
23.2%; Pred. No. 13
:ive 20; Mismatche
                                                                                                                                                                            NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESSEE: CURTIS, MORRIS & SAFFORD,
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
                                                                                                                                                                                                                                                                                                              NEW YORK
: UNITED STATES OF AMERICA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       123 GSKNTEV-----LVWEECVANSVVIL 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | :|:| :: | :: | :: | 482 GFYDTQVYDASPEIVSESMLVSVVAVI 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 25, Application US/09078173A
; Sequence 15, Application US/08473446; Patent No. 6017542; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/413,118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 533 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: peptide FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 48; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
```

```
496 DEDGSTCTMSEWI----TWSPCS-----VSCGMGMRSRERYVKQFPEDGSVCMLPTEETE 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 ----LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKT- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14 VYVNDSVWVPGPTDDRCPAKPEEEG-----MMINISIGYHYP------PICLGRAPGC- 60
                                                                                                       Sequence 10, Application US/07862021B
Patent No. 527996
GENERAL INFORMATION
FAPLICANT: Jessell, Thomas M
APPLICANT: Jessell, Thomas M
APPLICANT: Jessell, Thomas M
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 10, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
ITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROFEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,021B
FILING DATE: 19920405
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.9%; Score 66.5; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Pred. No. 32;
16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     116 -----CPKEIPKGSKNTEVLVWEECVA 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  547 KCTVNEECS --- PSSCLVTEWGEWDDCSA 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
NAME: White, John
REGISTRATION NUMBER: 26,678
REFERENCE/DOCKET NUMBER: 4002/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 664-0525
TELEFA: (212) 664-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                         E: Cooper & Dunham
30 Rockefeller Plaza
                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 23.5%;
Matches 35; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : 807 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; MOLECULE TYPE: protein US-07-862-021B-10
                                                                                                                                                                                                                                                                                                                              ADDA...
STREET: 30 ...
CITY: New YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                          10112
                                                                 RESULT 11
US-07-862-021B-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-313-288B-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH:
294 G 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |:||:| :|
177 LLPALQKFLFNFFSLTFLGADPSASPEIANSGFAYLDAWLAIQLAPTVSIGVLQPLEEIF 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98 LQDFS-----YQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEF 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98 LQDFS------YQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEF 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----PTVSP---NSRFTY-HMVSGMSLRPRVN-----Y 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35;
                                                                                                     APPLICANT: Alan Brash
APPLICANT: Nathalte Tijet
TITLE CANT: Nathalte Tijet
TITLE OF INVENTION: GUAVA (FSIDIUM GUAJAVA) 13-HYDROPEROXIDE
TITLE OF INVENTION: LYASE AND USES THEREOF
FILE REFERENCE: 06027.0001
CURRENT APPLICATION NUMBER: US/09/078,173A
CURRENT FILIOR DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FASTSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.9%; Score 66.5; DB 4; Length 480;
25.6%; Pred. No. 15;
Live 18; Mismatches 37; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 480;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Alan Brash
APPLICANT: Alan Brash
APPLICANT: Nathalie Tijet
TITLE OF INVENTION: WGSMELON (CUCUMIS MELO) HYDROPEROXIDE
TITLE OF INVENTION: LYASE AND USES THEREOF
FILE REFERENCE: 06027.0002
CURRENT APPLICATION NUMBER: US/09/537,357
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 32
LENGTH: 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 LMPAVQNWLVEV-----PTVSP---NSRFTY-HMVSGMSLRPRVN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 7.9%; Score 66.5; DB 4; Best Local Similarity 25.6%; Pred. No. 15; Matches 31; Conservative 18; Mismatches 37;
                                                                                                                                                                                                                                                                                                                                                                                       ) ORGANISM: Capsicum annum (green pepper) US-09-078-173A-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; ORGANISM: Capsicum annum (green pepper)
US-09-537-357-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 32, Application US/09537357
Patent No. 6271018
GENERAL INFORMATION:
         GENERAL INFORMATION:
APPLICANT: Ian M. Whitehead
APPLICANT: Alan Slusarenko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 25.65
Matches 31; Conservative
                                                                                      Duncan Gaskins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 LMPAVQNWLVEV --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-537-357-32
                                                                                                                                                                                                                                                                                                                    SEQ ID NO 25
LENGTH: 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              149 G 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            294 G 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        149 G 149
                                                                                 APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT
```

g

å

a

ô

8

ô

g

å

ð

ŝ

45;

Length 807; Indels

```
2252 WIDGSTLOYEN---WRPNQPDSFFSAGEDCVVIIWHENGQWNDVPCNYHLTYTCKKGTVA 2308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   496 DEDGSTCTMSEWI----TWSPCS-----VSCGMGMRSRERYVKQFPEDGSVCMLPTEETE 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            444 IYSNWSPW-----SACSSSTCEKGKRMRQRMLKAQLDLSVPCPDTQDFQPCMG--PGCS 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 ----LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKT- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14 VYVNDSVWVPGPTDDRCPAKPEEEG-----MMINISIGYHYP-----PICLGRAPGC- 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7 WMDNPIEVYVNDSVWVPGPTDDRCPAKPE-----EEGMMINISIGYHYPPICLGRAPG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       60 CLMPAVQNWLVE-----PRVSPISRFTYHMVSGMSLR------PRVN 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 66.5; DB 6; Length 2409;
Pred. No. 1.5e+02;
9; Mismatches 56; Indels 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7.9%; Score 66.5; DB 5; Length 807;
23.5%; Pred. No. 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Patent No. 5180808

; Patent No. 5180808;

; TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID

; SEQUENCES AND METHODS OF DETECTING THE SAME

; ANTIBODIES, AND METHODS OF DETECTING THE SAME

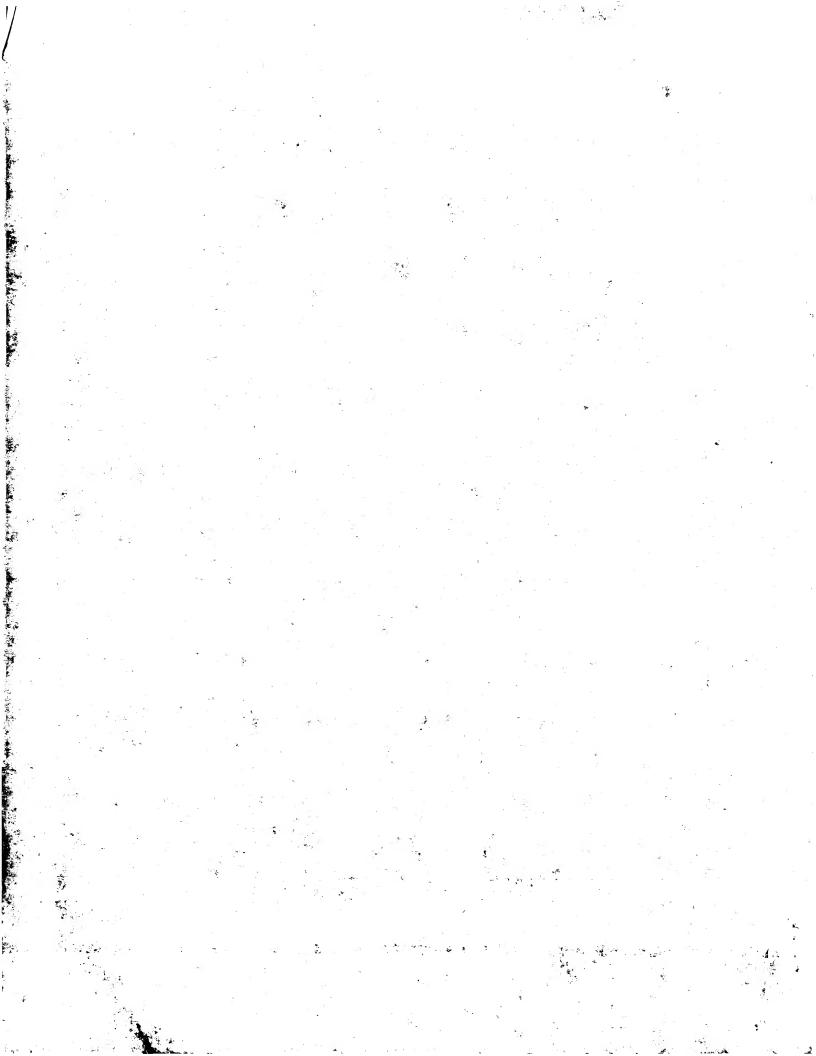
; NUMBER OF SEQUENCES: 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         116 -----CPKEIPKGSKNTEVLVWEECVA 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      547 KCTVNEECS---PSSCLVTEWGEWDDCSA 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/441,179
FILING DATE: 27-NOV-1989
                                                                                     ATTORNEY AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 664-0525
TELEX: 42523 COOP UI
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 807 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.9%;
                                                                                                                                                                                                                                                                                                                       : 807 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 23.5%;
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 7.94
Best Local Similarity 23.09
Matches 31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : :||| | :|:
2365 CMNPSAYQRTYSMKY 2379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97 YLQDFSYQR--SLKF 109
                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: protein
                       PRIOR APPLICATION DATA: APPLICATION NUMBER: FILING DATE:
    CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; LENGTH: 2409
5180808-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
US-09-258-754-451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NO:2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5180808-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKT- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             496 DEDGSTCTMSEWI----TWSPCS-----VSCGMGMRSRERYVKQFPEDGSVCMLPTEETE 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14 VYVNDSVWVPGPTDDRCPAKPEEEG-----MMINISIGYHYP------PICLGRAPGC- 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 807;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Albu Abblicant: Albu APPLICANT: Abblicant: Albu APPLICANT: Klar, Avihu
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham STREET: 30 Rockefeller Plaza
                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: TBM FO Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03164
FILING DATE: 19930402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEFROM: (212) 278-0400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 7.9%; Score 66.5; D
Best Local Similarity 23.5%; Pred. No. 32;
Matches 35; Conservative 16; Mismatches
2: Cooper & Dunham LLP
1185 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       547 KCTVNEECS---PSSCLVTEWGEWDDCSA 572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          116 -----CPKEIPKGSKNTEVLVWEECVA 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PCT-US93-03164-10; Sequence 10, Application PC/TUS9303164; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        807 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; MOLECULE TYPE: protein US-08-313-288B-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                            New York
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: New York
STATE: New York
COUNTRY: USA
                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10112
                                                                                                               10036
  ADDRESSEE:
                                            CITY: Nev
STATE: NE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

9

```
; Sequence 451, Application US/09258754
; Patent No. 6174687
; GRNERAL INFORMATION:
    APPLICANT: Rubalatini, Errki
    APPLICANT: Rajotte, Dailli, Relata
    APLICANT: Rajotte, Dailli, Relata
    CURRENT FILING DAIR: 1999-02-26
    CURRENT PILING DAIR: 1999-02-26
    SARLIER PILING DAIR: 1998-03-13
    CURRENT PILING DAIR: 1998-03-13
    CURRENT PILING DAIR: 1998-03-13
    SOFTWARR: Patentin Ver. 2.0
    TYPE: PRT
    ORGANISM: Mus musculus
    US-09-258-754-451
    Dail EECHMINISIG-----THYPPICLGRAPCIMPAVONNLV------EVPTVSP-NSRT 82
    Dail Haches 34; Conservative 15; Mismatches 56; Indels 27; Gaps 36
    ALT ECHLIDSCLOVIRTITHICANTVILLINGMRYLTHINGNRPWADNMLVDRODDEAESHGESPROKKEL 200
    SATH MATHORNING STORM STORM SATH SATHER SAYSLCPHRRNVPDOVLQLVKNTS 260
    Dail NEWTRICVAINGVAILESHVSVAIMEDALGISRAPVIFSHSSAYSLCPHRRNVPDOVLQLVKNTS 260
    Dail NEWTRICVAINGVAILO
```

Search completed: April 9, 2002, 16:57:24 Job time: 243 sec

261 SLVMVNFFSNFV 272



```
mus musculu
schizosacch
mus musculu
                           aquifex aeo
homo sapien
                                                    mus musculu
schizosacch
                                                                            mus musculu
  escherich1a
caenorhabd1
                                                                                           homo sapien
                                                                                                       drosophila
                                                                                                                                                                                                                                                                                                                                                                                                 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 AVQNWLVEVPTVSPISRFTYHWVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEIPKE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 AVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMP 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 PVTWMDNPIEVYVNDSIWVPGPIDDRCPAKPEEEGMMINISIGYRYPPICLGRAPGCLMP 60
                                                                                                                                                                                                                                                                                                                                                Eukaryota: Metazoa: Chordata; Craniata: Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID=9606;
P14081
Q21926
O66469
P53539
P13346
O94616
P35582
                                                                                                    P04146
Q62059
O14021
O89029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 584;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 768.5; DB 1;
Pred. No. 2.3e-69;
______ 6;
                                                                                                                                                                                                                                                                        01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
01-MAR-1989 (Rel. 10, Last annotation update)
RETROVIRUS-RELATED ENV POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SYI_CAEEL
SPSA_AOUNE
FOSB_MOUNE
FOSB_MOUNE
HNJA_MOUNE
PCHI_UNWA
COPI_DROME
GOCY_MOUNE
YDPI_SCHPO
MTN4_MOUNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91.5%;
illarity 94.7%;
Conservative
                                                                                                                                                                                                                                                                                                                                           sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
~~~~~~~~~~
                                                                                                                                                                                                                                               ENV1_HUMAN
P10267;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RAI2_MOUSE
Q9QVY8;
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RAIZ_MOUSE
ID RAIZ_MOU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                    ENV1_HUMAN
                                                                                                                                                                                                                                                                                                                                          Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         124
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                   (without alignments)
236.897 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             human papil
mus musculu
pichia angu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           homo sapien
rattus norv
rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amsacta moo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     homo sapien
schizosacch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             bacillus su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     saccharomyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   t amylopull
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           glycine max
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  t amylopull
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mouse mamma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  schizosacch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    schizosacch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 caenorhabdi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rhizobium
                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                      US-09-490-700-36
840
1 MVTPVTWMDNPIEVYVNDSV......ECVANSVVILQNNEFGTIID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                    April 9, 2002, 16:57:36; Search time 23.68 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 099/5p3

P26955

0639212

0639212

0639212

0639212

0639212

0639212

0639212

0639212

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

0639267

063927

063927

063927

063927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    043010
P32653
P31428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P10267
         GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                        hits satisfying chosen parameters:
                                                                                                                                                                                                                                 100059 segs, 36664827 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RAIZ_HUMAN
CYRB_MOUSE
S24A_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VL2_HPV54
P2CD_MOUSE
MASY_PICAN
ABOU_THEET
TREF_ECOLI
COMA_BACSU
NODU_RHITR
ENV_MMTVG
RPB2_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HN3A_HUMAN
PGH1_MOUSE
YG29_YEAST
FSPO_RAT
YQ36_CAEEL
MSH3_MOUSE
PGCV_HUMAN
                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENV_MMTVB
GLC3_SOYBN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AMEPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MRP_STRSU
MDP1_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HN3A_RAT
PGH1_RAT
APU_THETY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SCN1_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CTB2_MOUSE
                                                             protein search, using sw model
                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                   seq length: 0 seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                             SwissProt_39:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           470
598
1481
1481
549
214
575
688
1210
445
473
602
602
602
602
                                                                                                                                                                                                                                                           οį
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    66.5
66.5
66.5
66.5
66
                                                                                                                                       Title:
Perfect score:
                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                           Potal number
                                                                                                                                                                                                                                                                                   0B
0B
                                                             protein
                                                                                                                                                                   Sequence:
                                                                                                                                                                                                                                 Searched:
                                                                                                                                                                                                                                                                                                                                                                            Database
                                                                                                                                                                                                                                                                                   Minimum
Maximum
                                                                                     Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
```

;

N

```
SEQUENCE
                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                             CYRB_MOUSE
                                                                                                                                                                                                                                                                                                                                                                             QQ
  SSSSSSE
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-99168896; PubMed-10049581; Walpole S.M., Hiriyana K.T., Nicolaou A., Bingham E.L., Durham J., Vaulpole S.M., Hiriyana K.T., Nicolaou A., Bingham E.L., Durham J., Vauloula M., Sieving P.A., Trump D.; Identification and characterization of the human homologue (RAI2) of a mouse retinoic acid-induced gene in Xp22."; Genomics 55:275-283(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 LMPAVQNWLVEVPTVSPNSRFTYHMVS-GMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90 MPIHMQVEGSSAPELNPNGNATYVMTTQGPVQLPVVLEQHVFQHLNS-----PLVLPQE 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32 LITTEAWNINSTDLVKKALVTVPGPSILNPPAE-SQSGMALKVAATV-LQPLCLGESPVV 89
                                                                                                                                                                    Wissink S., Kruijer W.;
retinoic-acid-induced
                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25;
                                                                                                                                                                                                                                                                                                                                                                                                                                            10.7%; Score 89.5; DB 1; Length 529; 24.8%; Pred. No. 0.17; ive 22; Mismatches 65; Indels 2
                                                                                                                                                                                                                                                                                                                                                                                                     529 AA; 57178 MW; 10AA48B170FCDBD0 CRC64;
              20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
RETINDIC ACID-INDUCED PROTEIN 2 (3F8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
RETINOIC ACID-INDUCED PROTEIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           530 AA
                                                                                                                                                   MEDINE=94148137; PubMed=8314004;
Jonk L.J.C., de Jonge M.E., Vervaart J.M.,
"Isolation and developmental expression of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 IPKGSKNTEVLVWEECVANSVVILQNNEF 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       144 AP------CSSNAI---HNNLF 156
                                                                                                                                                                                                                                                                                                                                                                                          PRO-RICH
                                                                                                                                                                                                                                                                                                                                                             EMBL; X76652; -; NOT_ANNOTATED_CDS. MGD; MGI:1344378; Rai2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                           Dev. Biol. 161:604-614(1994).
-!- INDUCTION: BY RETINOIC ACID.
    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 10.77
Best Local Similarity 24.8'
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
20-AUG-2001 (Rel. 40, 20-AUG-2001 (Rel. 40, 20-AUG-2001 (Rel. 40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sapiens (Human).
                                                                    Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                               NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RAIZ_HUMAN
Q9Y5P3;
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                 genes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ношо
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                    ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ::| | :: | | :: | | | :: | | 32 LITTEAWNINSTDLVKKALVTVPAPSILNPPAE-SQSGMALKVAATV-LQPLCLGESPVV 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 LMPAVONWLVEVPTVSPNSRFTYHMVS-GMSLRPRVNYLQDFSYQRSLKFRPKGKTC 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      90 MPIHMQVEGSSAPELNPNGNATYVMTTQGPVQLPVVLEQHVFQHLNSPLVLPQEAPC 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                          9.8%; Score 82.5; DB 1; Length 530; 25.6%; Pred. No. 0.85; artive 19; Mismatches 65; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR00299; CRIA.
InterPro; IPR00299; CRIA.
InterPro; IPR00177; FN III.
InterPro; IPR00177; FN III.
InterPro; IPR00177; FN III.
Pfam; PF00041; fn3; 2.
PROSTIE; PS01355; HEMATOPO_REC_S_F1; 1.
Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
                                                                                                                                                                                                                                   PRO-RICH.
9879EE869DC6188F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
37-TYCKINE RECEPTOR COMMON BETA CHAIN PRECURSOR.
CSF2RB OR CSFR2BI OR AIC2B OR IL3RBI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             896 AA
                                                                                                                   or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                253 PI
57148 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M34397; AAA37204.1; -.
                                                                                                                                                                         EMBL; AF136587; AAD33688.1;
                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 25.68;
Matches 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; A35782; A35782
                                                                                                                                                                                                                                                             530 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                200
                                                                                                                                                                                                   MIM; 300217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CYRB_MOUSE
```

us-09-490-700-36.rsp

m

```
416 440 ZINC FINGER-LIKE.
1078 AA; 118791 MW; 9A3214F52D28FD3F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00250; FOCK_head; 1.
PRINTS; PR00053; FORKHEAD.
SMART; SM00339; FH; 1.
PROSITE; PS000659; FORK_HEAD_1; 1.
PROSITE; PS000659; FORK_HEAD_2; 1.
PROSITE; PS50039; FORK_HEAD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-91032994; PubMed-2227418;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSFAC; T00371; -. InterPro; IPR001766; Fork_head.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               48774 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; X55955; CAA39418.1; -.
                                                                                                                    Similarity 26.1%;
                                                                                                                                             31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; A36674; A36674.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      169 ;
466 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE-Liver,
                                                                                                                                                                                                                                                                                                                                                                                                                                                           HN3A_RAT
P23512;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA_BIND
SEQUENCE
                           SEQUENCE
                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                             Matches
       SO
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                         7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Orci L., Paccaud J.-P.;
Sec24 proteins and sorting at the endoplasmic reticulum.";
"Sec24 proteins and sorting at the endoplasmic reticulum.";
"Sec24 proteins and sorting at the endoplasmic reticulum.";
J. Biol. Chem. 274:7383-7840(1999).

-I-FUNCTION: COMPONENT OF THE COPII COAT, THAT COVERS ER-DERIVED
THE GOLGI APPARATUS. COPII ACTS IN THE CYTOPLASMIC RETICULUM TO
THE GOLGI APPARATUS. COPII ACTS IN THE CYTOPLASMIC PROMOTE THE
TROM THE ENDOPLASMIC RETICULUM TO THE GOLGI COMPLEX.

-I-SUBUNIT: COPII IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
COMPLEX. THE SEC13/31 COMPLEX AND SARI.

-I-TISSUE SPECIFICITY: EXPRESSED IN FIBROBLASTS, HEPATOCYTES, AND
LYMPHOCYTES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROTEIN TRANSPORT PROTEIN SEC24A (SEC24-RELATED PROTEIN A) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   794 IPGEPREEVGPASPHPEGLLVLQQVGDYCFLP---GLGPGSLSPH----SKPPSPSLCS 845
                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 VPG-PIDDRCPAKPEEEGMMINISIG-YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNS 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -1- SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.
                                       EXTRACELLURA (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.

BY SIMILARITY.

N-LINKED (GLCNAC. ..) (POTENTIAL).

N-LINKED (GLCNAC. ..) (POTENTIAL).

N-LINKED (GLCNAC. ..) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Transport; Protein transport; Golgi stack; Endoplasmic reticulum; Multigene family. NoN_TER 1\ 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Metazoa, Chordata, Cranlata, Vertebrata, Euteleostomi,
Mammalla, Eutherla, Primates, Catarrhini, Hominidae, Homo
                        CYTOKINE RECEPTOR COMMON BETA CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pagano A., Letourneur F., García-Estefanía D., Carpentier J.-L.,
                                                                                                                                                                                                                                                                                                                                                                                                         26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | :| :|:| | | | | 846 E-TEDLVQDLSVKKFPYQPMPQAPAIQFFKSLKHQDYLSLPPWDNSQSGKVC 896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    80 RFTYHMVSGMSLR-----RGKTC 116
                                                                                                                                                                                                                                                                                                                                                        9.0%; Score 76; DB 1; Length 896; llarity 25.9%; Pred. No. 6.9; Conservative 18; Mismatches 39; Indels
                                                                                                                                                                                                                                                                                     8CE16EDFDC07A999 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1078 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-B-cell;
MEDLINE-99175155; PubMed-10075675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AJ131244; CAA10334.1; -.
                                                                                                                                                                                                                                                                                   99111 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 29; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S24A_HUMAN
095486;
                        CHAIN
DOMAIN
TRANSMEM
                                                                                                                                                                                  DISULFID
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                 DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S24A_HUMAN
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FAMER REPRESENTATION OF THE PROPERTY OF THE PR
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'n
                                                                                                                                                                                                           272 SVGYSYP-----SLPPGYQNTTPPGATGVPPSSLNYPSGPQAFTQTPLGANHLTTS 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                85 --- HW---- 85
                                                                                                                                                                                                                                                                                     86 VSGMSLRPR----VNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSV 140
                                                                                                                                                                                                                                                                                                                              323 MSGLSLQPEGLRVVNLLQERNMLPSTPLKPPVPNLHEDIQKLNCNPELF---RCTLTSI 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUCH AS AFP, ALBUMIN, TYROSINE AMINOTRANSFERASE, PEPCK, ETC.
INTERACTS WITH THE CIS-ACTING REGULATORY REGIONS OF THESE GENES.
-! SUBCELLULAR LOCATION: NUCLEAR.
-! TISSUE SPECIFICITY: LIVER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA-binding; Nuclear protein; Transcription regulation; Activator. DOMAIN 251 288 ESSENTIAL FOR DNA BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally."; Genes Dev. 4:1427-1436(1990).
-I- FUNCTION: TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LIVER GENES
                                                                         34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16;
   Length 1078;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8.8%; Score 74; DB 1; Length 466; 25.2%; Pred. No. 5.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57; Indels
                                                                      41; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lai E., Prezioso V.R., Smith E., Litvin O., Costa R.H., Darnell J.E. Jr.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A., AND SEQUENCE OF 313-337 AND 413-434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06555BA74BD9B7DC CRC64;
                                                                                                                                            44 SIGYHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTY----
8.9%; Score 75; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1991 (Rel. 20, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HEPATOCYTE NUCLEAR FACTOR 3-ALPHA (HNF-3A).
HNF3A OR TCF3A OR TCF-3A.
                                                                      13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13; Mismatches
                                    Pred. No. 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FORK - HEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 25.2% Matches 29; Conservative
```

```
qq
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                    PGH1_RAT STANDARD; PRT; 602 AA.
063321; 062584;
15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROSTAGLANDIN G/H SYNTHASE 1 PRECURSOR (EC 1.14.99.1) (CYCLOOXYGENASE -1) (COX-1) (PROSTAGLANDIN-ENDOPEROXIDE SYNTHASE 1) (PGH SYNTHASE 1) (PGHS-1) (PHS 1).
PTGS1 OR COX1 OR COX-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CATALYTIC ACTIVITY: ARACHIDONATE + AH(2) + 2 O(2) = PROSTAGLANDIN
                                                                             PGPAASPQTLDHSGATATGGGSELKSPASSSAPPISSGPGGWICTPLSPTWLAPHESQLH 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Feng L., Sun W., Xia Y., Tang W.W., Chanmugam P., Soyooza E., Wilson C.B., Hwang D.; "Cloning two isoforms of rat cyclooxygenase: differential regulation of their expression.";
                       71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eling T.E.;
TPA-inducible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELL PROLIFERATION IN SOME NORMAL AND NEOPLASTICALLY TRANSFORMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-FISCHER 344; TISSUE-Trachea;
MEDLINE-95186876; PubMed-78644;
Kitzler J., Hill E., Hardman R., Reddy N., Philpot R., Eling T.E.;
"Analysis and quantitation of splicing variants of the TPA-inducible PGHS-I mRNA in rat tracheal epithelial cells.";
Arch. Blochem. Blophys. 316:856-863(1995).
Arch. Elochem. Ar PLAY AN IMPORTANT ROLE IN REGULATING OR PROMOTING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBUNIT: HOMODIMER (BY SIMILARITY).
SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MICROSOMAL MEMBRANE.
MISCELLANEOUS: THIS ENZYME ACTS BOTH AS A DIOXYGENASE AND AS A
                       PGP----TDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNWLVE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MISCELLANEOUS: THIS ENZYME IS THE TARGET OF NONSTEROIDAL ANTI-INFLANMATORY DRUGS SUCH AS ASPIRIN.
SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
SIMILARITY: BELONGS TO THE PROSTAGLANDIN G/H SYNTHASE FAMILY.
                                                                                                                                                                      382 LKGAPHYSFNHPFSINNL--MSSSEQQHKLDFKAYEQALQYSPYGATLPASLPLG 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         H2 + A + H(2)O.
PATHWAY: FIRST STEP IN THE FORMATION OF PROSTAGLANDINS AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN-SPRAGUE-DAWLEY;
MEDLINE-94099619; Pubmed-8274023;
Feng L., Sun W., Xia Y., Tang W.W., Chanmugam P., Soyoola E.,
                                                                                                                                        ---VPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Arch. Biochem. Biophys. 307:361-368(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002007; Anim_peroxdse.
InterPro; IPR000561; EGF-like.
InterPro; IPR001536; Peroxdse_3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001536; Peroxdse_3. PRINTS; PR00457; ANPEROXIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; U03388; AAA03465.1; -. EMBL; S67721; AAB29400.2; -. EMBL; U18060; AAA85823.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THROMBOXANES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PEROXIDASE.
                                                                                322
                          23
                                                                                                                                        72
                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                             HIDDRY REPAYED BY A COURT COUR
                                                                                g
                                                                                                                                        ŏ
```

```
WLRS--SLRPSPSFTHFLLTHGYWIWEFVNATFIREVLMRLVITVRSNLIPSPPTYNTAH 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                       DISTALL HESTIDINE (BY SIMILARITY).
CYCLOOXYGENASE (BY SIMILARITY).
PROXIMAL HEME LIGAND (BY SIMILARITY).
PROXIMAL HEME LIGAND (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 PPPVLLTDA-GVPSPVNPCCYYPCQNQGVCVRFGLD-HYQCDCTRTGYSGPNCTIPEIWT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLG---RAPGCLMPAVQN 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WLVEVPTVSPNSRFTYHM-------VSGMSLR--------PRVNYLQ 99
PROSITE; PSOU2; EGF_1; FALSE_NEG.
PROSITE; PSO1186; EGF_2; FALSE_NEG.
Oxidoreductase; Dioxygenase; Peroxidase; Glycoprotein; Acetylation;
Prostaglandin biosynthesis; Heme; Iron; Signal; Membrane;
EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       UL-AUG-1990 (Rel. 15, Last sequence update)
20-AUG-2001 (Rel. 40, Last sequence update)
AMYLOPULULANASE PRECURSOR (ALPHA-AMYLASE/PULLULANASE) [INCLUDES:
ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE);
(ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 602;
                                                                                               BY SIMILARITY.
PROSTAGLANDIN G/H SYNTHASE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB29400).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OBFAF7EBFDA0C7C2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Thermoanaerobacter thermohydrosulfuricus (Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DF-----SYQRSLKFRPKGKTCPKEIPKGSKNTEVL 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                thermohydrosulfuricum).
Bacteria, Firmicutes; Bacillus/Clostridium group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Thermoanaerobacter group; Thermoanaerobacter.
NCBL_TaxID=1516;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 72.5; DB Pred. No. 9.9; 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1475 AA
                                                                                                                                                                                                                                                                                                                                                                                     8.6%; Suc.
21.7%; Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      69032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                       119
192
263
274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      602 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
nes 35; Conserv
                                                                                                                                                                                                                                                                        571
70
70
106
146
412
36
1119
1192
263
274
                                                                                                              27
34
209
387
390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [1]
SEQUENCE FROM N
STRAIN=E101-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APU_THETY
P16950;
                                                                                                                                              ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                              CHAIN
                                                                                                                                                                                             MOD_RES
                                                                                                                                                                                BINDING
                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APU_THETY

APU_THETY

AC 01-80

DT 01-80

DT 01-80

DE AMPIC

DE ALPHE

DE PULLC

DE (ALPHE

GN APU.

OS THETY

OS THETY

OX NCBI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             137
```

6

S

```
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOYBN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                           DDT TO DD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1080 NAVYNKVEGVEISKDWTWVDAQYVGDSGNNDKYMAKFVPDMVGTWEYIMRFSSNQGHDWT 1139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1031 THVIGVNNPVEVYA--EVWAQGLTD-----KPGQGENMI-AQLGYRYIGDTVGDA---VY 1079
                                                                                                                                                                                                          Melasniem1 H.; "Purification and some properties of the extracellular alpha-amylase-pullulanase produced by Clostridium thermohydrosulfuricum."; Blochem. J. 250:813-818(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63 PAVQN------WL-VEVPTVSP---NSRFTYHMVSGM-----SLRPRVNYLQDFS 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 TPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLM 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                       Bork P., Doolittle R.F.;
"Proposed acquisttion of an animal protein domain by bacteria.";
Proc. Natl. Acad. Sci. U.S.A. 89.8990-8994(1992).
--- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART; SW00060; FN3; 2.
Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Repeat; Multifunctional enzyme.
                 Melasniem! H., Paloheimo M., Hemioe L.;
"Nucleotide sequence of the alpha-amylase-pullulanase gene from
Clostridium thermohydrosulfuricum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYZES (1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO FORM MALTOTRIOSE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1; Length 1475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3476C414110BE376 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FIBRONECTIN TYPE-III 1
FIBRONECTIN TYPE-III 2
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AMYLOPULLULANASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 27;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.6%; Score 72.5;
                                                                                           Gen. Microbiol. 136:447-454(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M28471; AAA23205.1; -.
PIK; A44765, A44765.
InterPro; IPR000461; Alpha_amylase.
InterPro; IPR001777; FN_III.
                                                                                                                                                                                                                                                                                                                                                      MEDLINE-93028390; PubMed-1409594;
                                                                                                                                                                                   MEDLINE-88268757; PubMed-3260488;
MEDLINE-90362027; PubMed-2391488;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00128; alpha-amylase; 1. Pfam; PF00041; fn3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14;
                                                                                                                                                                                                                                                                                                          [3]
FIBRONECTIN TYPE III DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         165631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similaring
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YTKG----PDGKT 1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1475
1016
1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    103 YORSLKFRPKGKT 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1163 12
629 65
658 65
735 73
                                                                                                                                          SECUENCE OF 32-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1140
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

RESULT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Glycine max (Soybean).

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 PVTW-MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYP-----PICL-- 54
                                   SEQUENCE FROM N.A.

MEDLINE-87112944; PubMed-3027377;

MODOR R., Dixon M., Smith R., Peters G., Dickson C.;

MOOLE R., Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshiff suppression events are required for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50;
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (Rel. 12, Created)
(Rel. 12, Last sequence update)
(Rel. 40, Last annotation update)
PRECURSOR (CONTAINS: GLYCININ A SUBUNIT; GLYCININ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8.6%; Score 72; DB 1; Length 688; 29.1%; Pred. No. 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31; Indels
                                                                                                                                            Mouse Manminary Lumor Virus (Strain Bro).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               46CAAC8D61FFADC2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LEADER PEPTIDE.
COAT PROTEIN GP52.
COAT PROTEIN GP36.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           481 AA
   688 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10; Mismatches
                                                                                                                                    (strain BR6)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                164 SFSTPTGCFQVDKQVFLSDTPTVDNN 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55 -- GRAPGCLMPAVQNWLVEVPTVSPN 78
                                                                                                                                                                                                                                                                                                     translation of gag and pol.";
J. Virol. 61:480-490(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-CV. DARE; TISSUE-Leaf;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M15122; AAA46544.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein; Polyprotein.
                                                                                                                                    Mouse mammary tumor virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             474
688
127
143
297
498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; D26795; VCMVMM.
Coat protein; Polypr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      498
557
688 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                      NCBI_TaxID=11758;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=3847;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-1989 (
20-AUG-2001 (
GLYCININ G3 P
SUBUNIT].
GY3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLC3_SOYBN
P11828;
ENV_MMTVB
P10259;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PM5P_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   482
  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the burspean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83 SYTNAPQEIYIQQGSGIFGMIFPGCPSTFEEPQQKGQSSRPQDRHQKIYHFREGDLIAVP 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45 IGYHYPPICLGRAPGCLMPAVQNWLV---EVPTVSPNSRFTYHMVSGMSLRPRVNYLQDF 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 143 TGFAY-------WMYNNEDTPVVAVSLIDTNSFQNQLDQMPRRFYLAGN 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: BELONGS TO THE 11S SEED STORAGE PROTEIN (GLOBULINS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Amsacta moorei entomopoxvirus (AmEPV).
Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
Entomopoxvirus B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         102 SYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTII 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1; Length 481;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INTERCHAIN (BY SIMILARITY 5F3C3148DF6241A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; S04605; S04605.
InterPro: IRR000459; Seedstore_11s.
Pfam; PF00190; Seedstore_11s; 1.
PRINTS; PR00439; 11SGLOBULIN.
PROSITE; PS00305; 11S_SEED_STORAGE; 1.
Seed storage protein; Multigene family; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (Rel. 25, Created)
(Rel. 25, Last sequence update)
(Rel. 33, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8.5%; Score 71.5; DI 18.1%; Pred. No. 9.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1002 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A SUBUNIT.
B SUBUNIT.
                  Cho T.-J., Nielsen N.C.;
"The glycinin Gy3 gene from soybean.";
Nucleic Acids Res. 17:4388-4388(1989).
MEDLINE-89296500; PubMed-2740231;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54241 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X15123; CAA33217.1; -. PIR; S04605; S04605.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   476
481
303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         481 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                       DISULFIDE BOND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=28321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1993
01-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SPHR_AMEPV
P29815;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SPHEROIDIN.
G5.
                                                                                                                                                                                                                                                                                                              FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
SPHR_AMEPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RREAR RRANGE RREAR RRANGE RREAR RRANGE RREAR RRANGE RREAR RRANGE RANGE RRANGE RRANGE RRANGE RRANGE RRANGE RRANGE RRANGE RRANGE RRAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Templeton N.S., Rodgers L.A., Levy A.T., Ting K.-L., Krutzsch H.C., Liotta L.A., Stetler-Stevenson W.G.;
"Cloning and characterization of a novel human cDNA that has DNA similarity to the conserved region of the collagenase gene family."; Genomics 12:175-176(1992).
                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE-92046310; PubMed-1942245;
Hall R.L., Moyer R.W.;
"Identification, cloning, and sequencing of a fragment of Amsacta moorei entomopoxvirus DNA containing the spheroidin gene and three vacchia virus-related open reading frames.";
J. Virol. 65:6516-6527(1991).
                     Banville M., Dumas F., Trifiro S., Arif B., Richardson C.; "The predicted amino acid sequence of the spheroidin protein from Amasacta moorei entomopoxvirus: lack of homology between major occlusion body proteins of different poxviruses."; J. Gen. Virol. 73:559-566(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -i - FUNCTION: MAJOR COMPONENT OF VIRAL OCCLUSION BODIES, THE PROTECTIVE COMPLEXES IN WHICH THE VIRIONS ARE EMBEDDED IN THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 1 ACETYLATION.
1002 AA; 114738 MW; DC0D61B157645F85 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CYTOPLASM OF THEIR INSECT HOSTS.
SUBUNIT: MAY FORM DISULFIDE BONDS LINKED AGGREGATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     440 CLKPKVPKNLRLWGWILDCDT---SRFIKHMADGSDDLDLDVRLN---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     - BINDICLKQAIKQHYTNVIILEYANTYPNCTLSLGNNRFNNVFD 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; JQ1436; PYVZAM.

Viral occlusion body; Acetylation; Late protein.

INIT_MET

MOD_RET

SEQUENCE 1002 AA; 114738 MW; DC0D61B157645F8!
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROTEIN PMS PRECURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8.5%; Score 71.5; Di
25.0%; Pred. No. 22;
iive 12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=92128984; PubMed=1310294;
MEDLINE=92185464; PubMed=1545219;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chordata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [1]
SEQUENCE OF 33-1222 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M75889; AAA42378.1; -. EMBL; M77182; AAA42383.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1-1148 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best_Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE=Melanoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PMSP_HUMAN
Q15155; P78421;
```

5

~

```
SPAC688.13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Simi
Matches 30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VL2_HPV54
Q81023;
                                                                                                                                                                                                                                                                                                                                                  Submitted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
VL2_HPV54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAC DDT TAN DD
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Loftus B.J., Kim U.J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D., "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q"; Genomics 60:295-308(1999).
                                                                                                                                                 -1- TISSUE SPECIFICITY: EXPRESSED IN COLON TUMOR TISSUE AND IN ADJACENT NORMAL COLONIC MUCOSA.
-1- DEVELOPMENTAL STAGE: NO DIFFERENCE BETWEEN NORMAL COLONIC MUCOSA AND COLON TUMOR TISSUE IN MRNA EXPRESSION, WHEREAS THE PROFEIN IS EXPRESSED 1.5-FOLD MORE IN NORMAL COLONIC MUCOSA THAT IN COLON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               56 RAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQ-DFSYQRSLKFRPKGK 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 PVTW------MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLG 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. . .) (POTENTIAL).
(GLCNAC. . .) (POTENTIAL).
(GLCNAC. . .) (POTENTIAL).
(GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 1222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DAB48007C6158841 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.)
N-LINKED (GLCNAC.)
N-LINKED (GLCNAC.)
N-LINKED (GLCNAC.)
EDI -> RDI (IN REF. 2).
S -> R (IN REF. 2).
K -> N (IN REF. 2).
N -> D (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PATISCHPO STANDARD; PRT; 335 AA. P41890; Q9P6L3; 01-NOY-1995 (Rel. 32, Created) 20-NOG-2001 (Rel. 40, Last sequence update) COT9 INTERACTING PROTEIN SCN1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.5%; Score 71.5;
19.2%; Pred. No. 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROTEIN PM5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HENE
 MEDLINE-99425270; PubMed-10493829;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X57398; CAA40655.1; -. EMBL; U91318; AAC15783.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    134321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1151 LPEQDIAGGS 1160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         115 TCPKEIPKGS 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                TUMOR TISSUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         218
369
568
33
200
404
458
493
513
513
513
650
650
650
650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query ...
Best Local Similar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SCN1_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              844488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-slb.ch/announce/or send an email to license@lsb-slb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202 DGRWIPSKTKIRKMKKEEYENSLAEER--KHYPPKICLHSYSGSIEQISQFSAHKVPT-- 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPP-ICLGRAPGCLMPAVQNWLVEVPTVS 76
                                                                                                                                                                                                                           MEDLINE-95096177; PubMed-7798319;
Samejima I., Yanagida M.;
"Bypassing anaphase by fission yeast cut9 mutation: requirement of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 335;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-972;
Aert R., Robben J., Volckaert G., Wood V., Rajandream M.A.,
Barrell B.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               43; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Delius H.; Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (APR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human papillomavirus type 54.
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DIC41E4A9DA19CEF CRC64;
Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               77 PNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGK 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---EFYYSFSIGINSRYK-NFIQ-----TLKGVPDDK 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Rel. 36, Created)
(Rel. 36, Last sequence update)
(Rel. 36, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          470 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.4%; Score 70.5;
Similarity 30.6%; Pred. No. 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D31845; BAA06631.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                Cell Biol. 127:1655-1670(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -!- FUNCTION: INTERACTS WITH CUT9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01026; TatD_DNase; 1.
SEQUENCE 335 AA; 38459 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AL355632; CAB90779.1; - InterPro; IPR001130; UPF0006.
                                                                                                                                                                                                                                                                                                                                     sut9+ to initiate anaphase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-JUL-1998 (Rel. 36, Lat
15-JUL-1998 (Rel. 36, Lat
15-JUL-1998 (Rel. 36, Lat
MINOR CAPSID PROTEIN L2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                   Schizosaccharomyces.
                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Papillomavirus.
NCBI_TaxID=37113;
                                                                                                                              NCBI_TaxID=4896
```

65722 MW; 4DB70B5D48539435 CRC64;

598 AA;

SEQUENCE

ω

```
4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce.or send an email to license@isb-sib.ch).
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROTEIN PHOSPHATASE ZC DELTA ISOCOBM (EC 3.1.3.16) (PP2C-DELTA) (P53-INDUCED PROTEIN PHOSPHATASE 1) (PROTEIN PHOSPHATASE MAGNESIUM-
                                                                                                                                                                                                                                         233 QVPVQDPAFLQQPSSLITYDNPVYEG---NPDVTLHFEQPTIHNAPD---PAFMDIFALH 286
                                                                                                                                                                                 Gaps
                                                                                                                                                                                                            13 EVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQN-WLVE 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                              17;
                                                                                                                                                Length 470;
                                                                                                                                                                                42; Indels
                                                        InterPro; IPR000784; Late_L2.
Pfam: PF00513; late_Protein_L2; 1.
Cast protein: Late protein.
SEQUENCE 470 AA; 50492 MW; CA0BFB5D8D05E5DF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00332; PP2Cc; 1.
SMART; SM00331; PP2C_SIG; 1.
RNGSITE; PS01032; PP2C; 1.
Hydrolase; Magnesium; Manganese; Multigene family.
DOMAIN 15 371
PP2C-LIKE.
                                                                                                                                                                                                                                                                       72 VPTVSP-----NSRFTYHMVSGMSLRPRVNYLQDFS 102
                                                                                                                                                                                                                                                                                       DB 1;
                                                                                                                                                                                                                                                                                                                                                                            598 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- INDUCTION: BY P53.
-1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
                                                                                                                                                                              16; Mismatches
                                                                                                                                                Score 70.5; I
Pred. No. 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro: IPR000222; PP2C.
InterPro: IPR003589; PP2C_catalytic.
InterPro: IPR001932; PP2C_domain.
InterPro: IPR001588; PP2C_sig.
Pfam: PF00481; PP2C; 1.
                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                       (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF200464; AAF09251.1; -.
                                                                                                                                                  8.4%;
                                          EMBL; U37488; AAA79192.1;
                                                                                                                                                              Best Local Similarity 25.7
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MGD; MGI:1858214; Ppmld
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEPENDENT 1 DELTA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-129/SVJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PPMID OR WIP1
                                                                                                                                                                                                                                                                                                                                                                                                       20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                            P2CD_MOUSE
                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                               RESULT 15
P2CD_MOUSE
SSEREN
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
10;
                                                     89 MSLRPRVNYLQDFSYQ-----SKNTE 128
                                                                                                                            71; Gaps
                                           3 TPVTWMDNPIEVYV---NDSVW--VPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRA 57
                                                                              58 PG-----LVEVPTVSPNSRFTYHMVSG 88
      Length 598;
                          Indels
       DB 1;
                                                                                                                                                               129 VLVWEECV------ANSVVILQNNEFGTIID 153
8.4%; Score
22.9%; Pred. No. 10,
ive 23; Mismatches
                          Conservative
                Similarity
                        50;
       Query Match
Best Local S
                         Matches
                                                                                                                                   406
                                                                                                                                                                     466
                                                          op
                                                                                                                                 qq
                                                                              δ
                                                                                               q
                                                                                                                 ΩŽ
                                                                                                                                                   οy
```

Search completed: April 9, 2002, 17:07:01 Job time: 565 sec

```
April 9, 2002, 16:48:46; Search time 72.35 Seconds (without alignments) 156.644 Million cell updates/sec
                                                                                                                                                                                                                                            1 MVTPVIWMDNPIEVYVNDSV......ECVANSVVILQNNEFGTIID 153
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                       Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                    522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                              OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                              US-09-490-700-36
840
                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A_Geneseq_1101:*
                                                                                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                    Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database
                                                                                                                       Run on:
```

1. SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
4. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
5. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
6. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
7. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
8. SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
9. SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT:*
11. SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT:*
12. SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13. SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
14. SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
15. SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
16. SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
17. SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT:*
18. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
18. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
19. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
11. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
12. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
13. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
14. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
15. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
16. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
17. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
18. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
18. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*
22. SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	Human endogenous r	Peptide #486 encod	Peptide #495 encod	Peptide #474 encod	Recombinant alpha	Mouse mammary tumo					
	ID	AAW95692	AAW97745	AAW95694	AAW97747	AAW95693	AAW97746	AAM14052	AAM26458	AAM01792	AAR08221	AAY41142
	Query Match Length DB ID	153 20	153 20	181 20	181 20	561 20	561 20	48 22	48 22	48 22	1475 11	688 20
øР	Query Match Le	100.0	100.0	100.0	100.0	100.0	100.0	26.2	26.2	26.2	9.8	9.8
	Score	840	840	840	840	840	840	220	220	220	72.5	72
	Result No.	П	7	m	4	S	9	7	80	6	10	11

AmEPV Spheroidin page AmEPV entomopoxvir Spheroidin protein E. coli ironec ext Mouse membrane dip Human polypeptide Arabidopsis thalia Human secreted pro Human polypeptide Buman polypeptide Human colon cancer Feline herpes viru Protein involved i Human concer case of the Aman colon cancer Feline herpes viru Protein involved i	F-spondin (FP5-9). Rat vascular smoot Vascular endothell Versican. Homo sa Novel bone marrow Neisseria meningit Plasmodium falcipa Muramidase release Dehydropeptidase-1 Hepatitis C virus
AAR29648 AAW30169 AAW30169 AAX30169 AAX30162 AAW39062 AAG47392 AAG47355 AAG44355 AAG44355 AAG31811 AAG31811 AAG31811 AAG31810 AAW39273 AAW39273 AAW39273 AAW39006 AAW36006 AAW3600	AAR4241 AAY0264 AAB13074 AAR12609 AAY15590 AAY75591 AAY75591 AAX77591 AAX77591 AAX77591 AAX7746 AAR53920
1153 2022 2022 2022 2022 2022 2022 2022 20	
10003 10003 10003 10003 1252 252 2552 2556 2664 185 1965 1565 1565 1565 1565 1565 1765 1765 17	807 807 2409 488 704 704 1256 410 3023
00000000000000000000000000000000000000	
71.5 71.5 71.5 71.5 71.5 70.5 70.5 69 69 69 69 69 69 69 69 69 69 69 69 69	
25 4 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8

ALIGNMENTS

Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity – new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope. Human endogenous retrovirus IDDMK1.2-22 env protein. AAW95692 standard; Protein; 153 AA 97EP-0401773. 97EP-0401773. (first entry) WPI; 1999-097928/09. Conrad B, Mach B; (MACH/) MACH B F. N-PSDB; AAX07514. 23-JUL-1997; Homo sapiens 23-JUL-1997; 08-JUN-1999 EP893691-A1. 27-JAN-1999 AAW95692; RESULT AAW95692 THOUSE STATE OF THE STATE OF TH

```
8 × 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 × 6 4 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                   The retrovirus has Superantigen (549) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence mellitus. Modified proteins expressed by the retroviral sequence (without SA9 activity but still able to induce an immune response) are useful in vaccines to treat or prevent SA9-related autoimmune disease; uncleic acid sequences encoding (modified) SA9 can be used similarly to treat such diseases, Retroviral-encoded SA9 are important in pathogenesis of autoimmune disease, probably by activating autoimmune disease, probably by activating between expressed and non-expressed viral nucleic acids) and can be used even where the pathogen is an ubiquitous endogenous retrovirus. Blood corporate and the pathogen is an ubiquitous endogenous retrovirus. Blood diagnosis can be made before clinical signs are apparent, allowing early intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    proteins, antibodies and specific binding agents, used for treating
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HERV; IDDKK1.2-22; superantigen; SAg; antigen; IDDM;
insulin-dependent diabetes mellitus; autoimmune disease; diagnosis;
therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                The sequence is that of an insulin-dependent diabetes mellitus associated human endogenous retrovirus (IDDMK1.2-22) env protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human endogenous retrovirus IDDKK1.2-22 envelope protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 840; DB 20;
100.0%; Pred. No. 8.5e-88;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 pkgskntevlvweecvansvvilqnnefgtiid 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW97745 standard; Protein; 153 AA
                         preventing autoimmune disease
                                                                   Claim 31; Fig 7D; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human endogenous retrovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97EP-0401773
97EP-0112482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         al Similarity 100.
153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42..45
/note= '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9905527-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     04-FEB-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW97745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
AAW97745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
```

```
This is the envelope protein (Env) of a new human endogenous retrovirus (HERY), designated IDDMX1.2-22, that has been identified as the source of Superantigen (SAP) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERY encodes SAP activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV caping SAP activity comprises specifically detecting in a biological sample either: (a) the mRNA of an expressed HERV having SAP activity (especially IDDMX1.2-22 5'LTR, 3' long terminal repeat, env or pol); (b) a protein or peptide expressed by the HERV (see AAW97745-48); (c) antibodies specifically associated with the HERV. Products of the invention can be used to identify substances capable of blocking transcription or translation of SAP encoding nucleic acid sequences, useful in the SAP, and/or prevention of autoimmune disease associated with the SAP. A nucleic acid encoding the endogenous SAP in IDDM suggests a general model according to which self SAP-driven and systemic activation of autoreactive T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 mvtpvtwmdnpievyvndsvwvpgptddrcpakpeeegmminisigyhyppiclgrapgc 60
                                                                                                                                                                                                                                                         New isolated human endogenous retrovirus - used to develop products for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human endogenous retrovirus IDDMK1.2-22 env/fs (sag) protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 100.0%; Pred. No. 8.5e-88;
Matches 153; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 pkgskntevlvweecvansvvilgnnefgtiid 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AA.
                                                                                                                                                                                                                                                                                                                                                                                                                Claim 6; Fig 7D; 165pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW95694 standard; Protein; 181
                                                                                                                                                                                   N-PSDB; AAX07188, AAX07189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JUN-1999 (first entry)
                                                                                                                                                WPI: 1999-143118/12
(MEDI-) MEDIGEN SA.
                                                                         B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EP893691-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW95694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                         Conrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW95694
```

ö

m

```
Human endogenous retrovirus
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                       associated human endogenous retrovius (IDDMK1.2-22) env/fs protein.

The retrovius has Superantigen (SAg) activity. It can be used
as part of a method is specifically used to diagnose type I diabetes
mellitus. Modified proteins expressed by the retroviral sequence
(without SAg activity but still able to induce an immune response)
are useful in vaccines to treat or prevent SAg-related autoimmune
disease; nucleic acid sequences encoding (modified) SAg are important
in pathogenesis of autoimmune disease, probably by activating
autoreactive T cells. The method is very specific (it can differentiate
between expressed and non-expressed viral nucleic acids) and can be used
con where the pathogen is an ubiquitous endogenous retrovirus. Blood
cor plasma samples can be tested without extensive preparation and
diagnosis can be made before clinical signs are apparent, allowing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                          Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating or preventing autoimmune disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HERV; IDDKKI.2-22; superantigen; SAg; antigen; IDDM;
insulin-dependent diabetes mellitus; autoimmune disease; diagnosis;
therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .;
0
                                                                                                                                                                                                                                                                                                                                                             sequence is that of an insulin-dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human endogenous retrovirus IDDKK1.2-22 Env/F-S (SAg).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 840; DB 20;
llarity 100.0%; Pred. No. 1.1e-87;
Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAW97747 standard; Protein; 181 AA
                                                                                                                                                                                                                                                                                                                            Claim 31; Fig 7G; 92pp; English.
                                                                           97EP-0401773
                                            97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                             WPI; 1999-097928/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 153; Conserv
                                                                                                                                              Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AA;
                                                                                                            (MACH/) MACH B F.
                                                                                                                                                                                             N-PSDB; AAX07516
                                                                           23-JUL-1997;
                                            23-JUL-1997;
             27-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW97747;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                             The
```

a ð 合 ð g

```
This is the envelope (Env)/FS polypeptide of a new human endogenous retrovirus (HERV), designated IDDMK1.2-22, that has been identified as the source of superantigen (SAG) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAG activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autolimune disease associated with a HERV call ample either: (a) the mRNA of an expressed HERV having SAG activity (especially IDDMK1.2-22 5'LTR, 3' long terminal biological sample either: (a) the mRNA of an expressed by the HERV (see AAW97745-4B); (c) antibodies specifically associated with the HERV. Products of the invention can be used to identify substances capable of blocking transcription or translation of SAG-encoding nucleic acid sequences, useful in therapy and/or prevention of autolimmune disease associated with the SAG. A nucleic acid encoding the endogenous SAG in IDDM suggests a general model according to which self SAG-driven and systemic activation of autoreactive T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated human endogenous retrovirus – used to develop product: for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%; Score 840; DB 20; 100.0%; Pred. No. 1.1e-87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                     "Asn is N-glycosylated"
                                                                                                          /note- "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 6; Fig 7G; 165pp; English.
                                                                                                                                                                                                                                                                                                                                        97EP-0401773.
                                                                                                                                                                                                                                                                                   98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                          17..19
/note=
                                                                                  42..45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-143118/12
                                                                                                                                                                                                                                                                                                                                                                                                                           (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAX07191
                                                                               Modified-site
                                                                                                                                                                                                                                                                                22-JUL-1998;
                                                                                                                                                                  W09905527-A2
                                                                                                                                                                                                                                                                                                                                        23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                      22-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conrad B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
```

ö

```
The sequence is that of an insulin-dependent diabetes mellitus associated human endogenous retrovirus (IDDMK1.2-22) env/fs-sag protein. The retrovirus has Superantique (S49) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence (without S49 activity but still able to induce an immune response) are useful in vaccines to treat or prevent S49-related autoimmune disease, nucleic acid sequences encoding (modified) S49 can be used similarly to treat such disease, probably by activating in pathogenesis of autoimmune disease, probably by activating autoractive T cells. The method is very specific (it can differentiate between expressed and non-expressed viral nucleic acids) and can be used even where the pathogen is an ubiquitous endogenous retrovirus. Blood cor plasma samples can be tested without extensive preparation and diagnosis can be made before claincal signs are apparent, allowing carly intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating or preventing autoimmune disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 561;
                                                                                                                                                                                                                                                                                 Human endogenous retrovirus IDDMK1.2-22 env/fs-sag protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 840; DB 20;
100.0%; Pred. No. 4.6e-87;
ive 0; Mismatches 0;
                                                                 AAW95693 standard; Protein; 561 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 31; Fig 7E; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97EP-0401773
                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 100.
Matches 153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI: 1999-097928/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                561 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (MACH/) MACH B F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAX07515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-JUL-1997;
                                                                                                                                                                                                        08-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27-JAN-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP893691-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                       AAW95693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
RESULT
AAW95693
```

```
This is the envelope protein (Env) of a new human endogenous retrovirus (HERV), designated IDDMK1.2-22, that has been identified as the source of superantigen (SAg) activity in insulin-dependent diabetes mallitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAg activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV having SAg activity comprises specifically detecting in a biological sample either: (a) the mRNA of an expressed HERV having capetivity (especially IDDMK1.2-22 5'LTR, 3' long terminal repeat, env or pol); (b) a protein or peptide expressed by the HERV: or (d) SAg activity specifically associated with the HERV. See AAW97745-48); (c) antibodies specific to the proteins expressed the HERV. See AAW97745-48); (c) antibodies specifically associated with the HERV. Products of the invention can be used to identify substances capable of blocking transcription or translation of SAg-encoding contents.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         autoimmune disease associated with the SAg. A nucleic acid encoding human retroviral SAg can be used as a DNA vaccine. Expression of the endogenous SAg in IDDM suggests a general model according to which self SAg-driven and systemic activation of autoreactive T cells leads to organ-specific autoimmune disease.
                                                                                                                                                                                                                                            HERV; IDDKKI.2-22; superantigen; SAg; antigen; IDDM; insulin-dependent diabetes mellitus; autoimmune disease; diagnosis; therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New isolated human endogenous retrovirus - used to develop products for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                        Human endogenous retrovirus IDDKK1.2-22 envelope protein.
                                                                                                                                                                                                                                                                                                                                                                                   17..19
/note= "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "Asn is N-glycosylated"
121 pkgskntevlvweecvansvvilqnnefgtiid 153
                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                             Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 6; Fig 7E; 165pp; English.
                                                                                           AAW97746 standard; Protein; 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-EP04926.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97EP-0401773.
97EP-0112482.
                                                                                                                                                                                                                                                                                                                           Human endogenous retrovirus
                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1999-143118/12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          561 AA;
                                                                                                                                                                                                                                                                                                                                                                 Key
Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                    Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9905527-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22-JUL-1997;
                                                                                                                                                                    21-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-FEB-1999
                                                                                                                               AAW97746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                         AAW97746
                                                       RESULT
```

DB 20; Length 561;

100.0%; Score 840;

Query Match

LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120

61 121

g δ

ò Ω PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153

Matches 153;

ò 음 ö 61 121 121

임

8

ò

ö

Gaps

; 0

Indels

```
The present invention relates to single exon nucleic acid probes (SENP: see AAI31315-AAI57546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of
                                                                                                                                                                                                                                                               Peptide #495 encoded by probe for measuring placental gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         probes useful
                                microarray; human; placenta; antenatal diagnosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        47 YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPR
 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 220; DB 22;
Pred. No. 7.8e-18;
1; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human genome-derived single exon nucleic acid analyzing gene expression in human placenta -
 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 27; SEQ ID No 26727; 654pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rank DR;
                                                                                                                                                         AAM26458 standard; Protein; 48 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM01792 standard; Protein; 48 AA
 ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0234587.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26.2%;
83.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-JAN-2001; 2001WO-US00663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                              (first entry)
 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                human genetic disorders.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-488897/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 40; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  48 AA;
                                                                                                                                                                                                                                                                                                                   genetic disorder
                                                                                                                                                                                                                                                                                                                                                                                     WO200157272-A2.
                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          04-FEB-2000;
                                                                                                                                                                                                                            17-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                       09-AUG-2001
 40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM01792;
                                                                                                                                                                                            AAM26458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                   Probe;
                                                                                                                       œ
 Matches
                                                                                                                                         AAM26458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAM01792
                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                    δ
                                                                   g
                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X Y X E
                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   present invention relates to human single exon nucleic acid probes
                                                                                                                                       61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                  Gaps
                                                  1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGGMMINISIGYHYPPICLGRAPGC 60
                                                                    Peptide #486 encoded by probe for measuring cervical gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      human; microarray; gene expression; cervical epithelial cell;
                  .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Human genome-derived single exon nucleic acid probes useful f
analyzing gene expression in human cervical epithelial cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 48;
                  Indels
                ;
0
 Pred. No. 4.6e-87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 220; DB 22;
Pred. No. 7.8e-18;
                  Mismatches
                                                                                                                                                                                                         PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 27; SEQ ID No 18878; 487pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Chen W, Rank DR;
                                                                                                                                                                                                                                                                                                                   Ş
              ;
0
                                                                                                                                                                                                                                                                                                              AAM14052 standard; Protein; 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-UUN-2000; 2000US-0693408.
03-AUG-2000; 2000US-053366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0234687.
04-OCT-2000; 2000GB-0024263.
100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JAN-2001; 2001WO-US00670
                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2001-488901/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               À,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cervical cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200157278-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
```

09-AUG-2001

Penn SG,

Sequence

12-OCT-2001

AAM14052;

AAM14052

for

ö

Gaps

ö

Indels

Length 48;

```
P-PSDB; AAR08221
                                                                                                                                          05-JUN-1990;
                                                                                                                                                                       05-JUN-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-JAN-2000
                                                                                                          12-DEC-1990
                                                                          EP402092-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY41142;
              Key
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present invention relates to novel single exon nucleic acid probes (see AA100010-AA110067). The present sequence is a peptide encoded by one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probe hybridises at high stringency to a nucleic acid expressed in the human breast. The probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aetiology. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative
           Peptide #474 encoded by probe for measuring human breast gene expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              breast disease and non-carcinoma tumours.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                           Novel single exon nucleic acid probe used to measuring gene expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                          Probe; human; breast disease; breast cancer; development disorder;
inflammatory disease; proliferative breast disease; non-carcinoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Starch; pullulan; saccharification; debranching; liquefaction;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 220; DB 22;
Pred. No. 7.8e-18;
1; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Recombinant alpha amylase pullulanase enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Clostridium thermohydrosulphuricum DSM 3783.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 27; SEQ ID No 10532; 322pp; English
                                                                                                                                                                                                                                                                                                                                                                              Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR08221 standard; protein; 1475
                                                                                                                                                                                                                                                                                                                                               (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                   26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-06608408.
03-AUG-2000; 2000US-0663266.
21-SEP-2000; 2000US-0236575.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                              Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26.2%;
83.3%;
                                                                                                                                                                                      29-JAN-2001; 2001WO-US00661.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                              Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-476286/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
les 40; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                            a human breast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              48 AA;
                                                                                                                      WO200157270-A2.
                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-MAR-1991
                                                                                                                                                       09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hydrolysis
                                                                                                                                                                                                                                                                                                                                                                              Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR08221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR08221
```

qq

à

```
1031 thvigvnnpvevya--evwaggltd-----kpgggenmi-aqlgyryigdtvgda---vy 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The recombinant enzyme was produced by cloning the gene encoding it from a C. thermohydrosulphuricum genomic library and using it to express the enzyme in a host cell e.g. E.coli which is easier to cultivate and has less complex nutritional requirements than C. thermohydrosulphuricum (which is an obligate anaerobic thermophilic organism). The enzyme is used to hydrolyse starch, amylase or pullulan, opt. in conjunction with a glucogenic or maltogenic enzyme. It has an optimum temp. range of 80-85 deg.C, i.e. 5 deg. lower than that of the native enzyme, but the heat stability is the same. At least 10 differently sized polypeptides having the same enzymic activities were produced, of Mr 100 000 - 165 000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63 PAVQN------WL-VEVPTVSP---NSRFTYHMVSGM-----SLRPRVNYLQDFS 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 TPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLM 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Heat stable enzyme with both alpha-amylase and pullulanase activities - prepd. by expressing Clostridum thermohydrosulphuricum DNA in host cells, useful in hydrolysis of starch, etc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mouse mammary tumor virus; MMTV; immunomodulatory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB, 11; Length 1475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mouse mammary tumor virus (MMTV) env polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8.6%; Score 72.5; DE
28.6%; Pred. No. 44;
Live 14; Mismatches
                                                              /label= signal peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Immune response; mouse mammary tumor epitope; passive immunotherapy; env.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY41142 standard; Protein; 688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 6; Fig 5; ]27pp; English.
                                                                                                                                                                                                                                                                                    90EP-0306088
                                                                                                                                                                                                                                                                                                                                                          89US-0361368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Welasniemi H, Paloheimo M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mouse mammary tumor virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1140 ytkg----pdgkt 1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           103 YORSLKFRPKGKT 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1990-370103/50.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1475 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 38; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                            (ALKO-) ALKO LTD.
```

6

us-09-490-700-36.rag

```
The sequences given in AAR29646-51 are encoded by a DNA fragment derived from the Entomopoxvirus Amsacta moore! (AmEPV) around the spheroidin DNA sequence. The open reading frames encode the spheroidin protein itself and also other structural or requilatory genes associated with spheroidin. EPV spheroidin has no significant amino acid homology to any previously reported protein. It is a non-essential protein which exogenous DNA. The spheroidin gene is naturally expressed at high levels. Small fragments of the surrounding DNA can be used as regulatory sequences if placed in operative association with foreign
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                        New viral vectors and chimeric vaccines - comprise entomopoxvirus expression system contg. spheroidin or thymidine kinase sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New entomopoxvirus polynucleotide sequences, proteins and vectors - are used for expression of heterologous proteins in both insect
                                                                                                                                                                                                                                                                                                                                                                                                                           DB 13; Length 1003;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Spheroidin; gene expression; vector; insect cell culture; mammal cell culture; AmEPV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      37; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             441 clkpkvpknlrlwgwildcdt----srfikhmadgsddldldvrln----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Pred. No. 34;
12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                          8.5%; Score 71.5; 25.0%; Pred. No. 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR55576 standard; Protein; 1003 AA.
                                                                                                                                                                   Disclosure; Fig 2; 110pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Moyer RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Amsacta moorei entomopoxvirus.
                                             Gruidl ME, Hall RL, Moyer RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93WO-US11907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92US-0991867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       26; Conservative
                  (UYFL ) UNIV FLORIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (UYFL ) UNIV FLORIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1994-217887/26.
                                                                            WPI; 1992-316172/38
                                                                                                                                                                                                                                                                                                                                                                              1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AmEPV spheroidin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAQ66808.
                                                                                          N-PSDB; AAQ28301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            07-DEC-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-JAN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-DEC-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-JUN-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO9413812-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gruidl ME,
                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR55576;
                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR55576
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5
                                                                                                                                                                                                                                                                                           The invention provides a method for inducing an immune response to mouse mammary tumor virus (MMTV) in a human subject. The method comprises administering to the subject an immunomodulatory composition comprising a pharmaceutical carrier and at least one MMTV antigen (or a discontinuous specific cellular or humoral response. The compositions and methods of the invention are useful for treating or preventing a human disease or disorder caused by MMTV, or a biological or chemical agent that reacts with an antiboddy directed towards an MMTV antigen. The immunomodulatory exposed to MMTV or MMTV antigen. The immunomodulatory exposed to MMTV or MMTV antigen ractive agents. The immunomodulatory compositions can be used in a preventative manner for those subjects not have use in a diagnostic assays. The present sequence represents the
                                                                                                                                                                                                                          Novel immunomodulatory compositions of mouse mammary tumor virus (MMTV) antigens used for treating or preventing MMTV infections \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 PVTW-MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYP-----PICL-- 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Entomopoxvirus; non-essential; regulatory sequences; vector.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8.6%; Score 72; DB 20; Length 688; 29.1%; Pred. No. 18; tive 10; Mismatches 31; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid sequence of the MMTV env polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR29648 standard; Protein; 1003 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             55 -- GRAPGCLMPAVQNWLVEVPTVSPN 78
                                                                                                                     (ITBI-) INT BIOIMMUNE SYSTEMS INC.
                                                                                                                                                                                                                                                                      Disclosure; Fig 5; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AmEPV Spheroidin protein G5R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91US-0657584.
                                                           99WO-US07712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92WO-US00855
                                                                                       98US-0057003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 29.18
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                  Stewart THM, Gershoni J;
                                                                                                                                                                              WPI; 1999-620171/53.
N-PSDB; AAZ23193.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Amsacta moorei.
W09951268-A1
                                                         08-APR-1999;
                                                                                       08-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-FEB-1991;
30-JAN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03-SEP-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR29648;
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

ŝ

ω

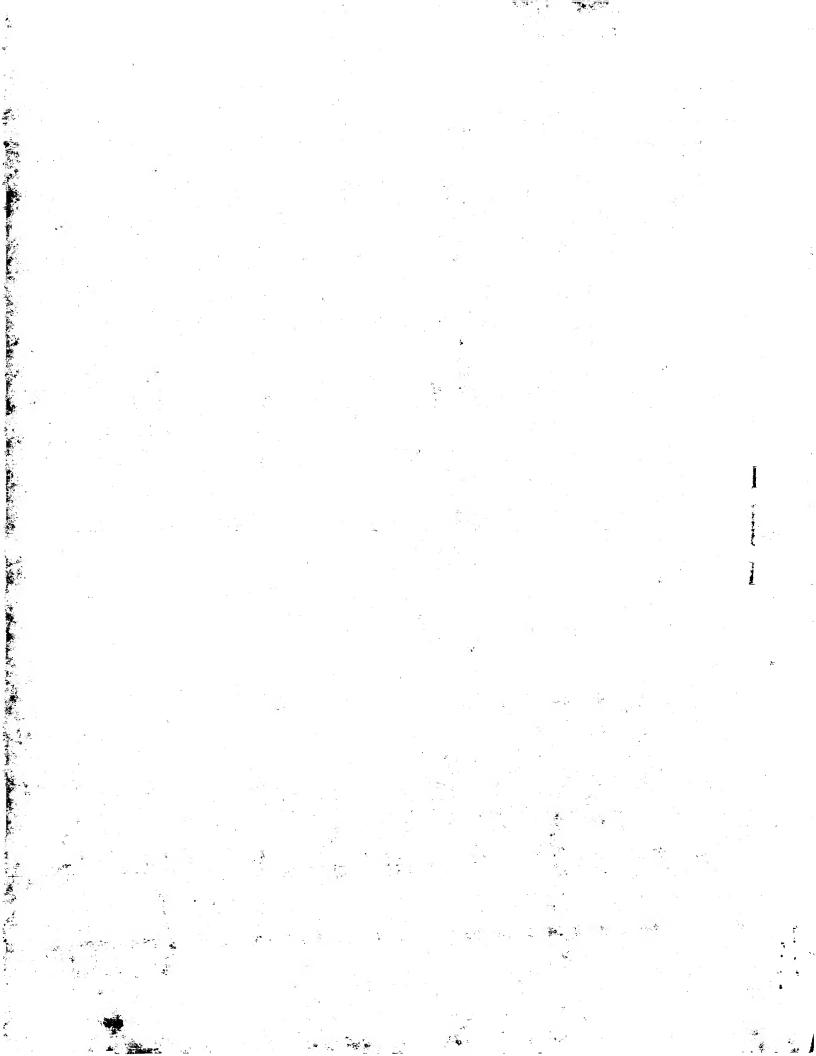
```
1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17-0CT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-FEB-1992;
07-DEC-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-FEB-1991;
30-JAN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JS5935777-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-ocr-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sruidl ME,
                                                                                                                                                                                                                                                                                                                                                                                                                      AAY30169;
    Sequence
                                                                                                                                                                                                                                                                                                                                                          AAY30169
  g
                                                                                                                                                                                                                                                                                                                                                                             THE STATE OF THE S
                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                        9
                                                                                                                                                   ŏ
                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           spheroidin gene, which is an example of the gene of the invention, which encodes a 115 kDa protein. EPV spheroidin and thymidine kinase promoters can be used in DNA constructs and vectors for expression of heterologous genes in insects or mammalian cells, e.g. vectors containing Bactilus thuringiensis toxin genes for use in insect control, or recombinant vaccinia or swinepox viruses for use as
                                                                                                                                                                                                                                                                                                                                    CLMPAVO-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This sequence is encoded by the Amsacta moorei entomopoxvirus (AmEPV)
                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Entomopoxvirus; spheroidin gene; AmEPV; thymidine kinase; promoter; insect control; viral vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New Entomopoxvirus nucleic acid sequences - used in DNA constructs and vectors for expression of heterologous genes in, e.g. insect
                                                                                The sequence of the Amsacta moorei entomopoxvirus spheroidin gene and its flanking regions was determined. The spheroidin gene can be used as the location for the insertion of heterologous DNA in insect and mammalian expression systems.
                                                                                                                                                                                                                                                                                           29;
                                                                                                                                                                                                                                                    DB 15; Length 1003;
                                                                                                                                                                                                                                                                                           37; Indels
                                                                                                                                                                                                                                                                                                                                                                         441 clkpkvpknlrlwgwildcdt----srfikhmadgsddldldvrln--
                                                                                                                                                                                                                                                                                                                                                                                                                                              111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AmEPV entomopoxvirus spheroidin protein G5R.
                                                                                                                                                                                                                                                  Query Match 8.5%; Score 71.5; C
Best Local Similarity 25.0%; Pred. No. 34;
Matches 26; Conservative 12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Columns 45-50; 55pp; English.
                                         Disclosure; Page 68-71; 118pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW41301 standard; Protein; 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (UYFL ) UNIV FLORIDA RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93US-0107755.
91US-0657584.
92US-0827685.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93US-0107755.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Moyer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92WO-US00855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAV14507, AAV14517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
and mammalian host cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gruidl ME, Hall RL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1998-168476/15
                                                                                                                                                                                       1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Amsacta moorei.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             viral vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22-NOV-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-NOV-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US5721352-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-FEB-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-JAN-1992;
12-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-FEB-1998
                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW41301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                    9
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                    õ
```

```
AAY30165-70 represent proteins encoded by open reading frames (ORFs) of the Amsacta moorei entomopoxvirus spheroidin gene and flanking sequences. The DNA is used to make expression systems of the invention. The specification describes an Entomopoxvirus (EPV) expression system that is capable of directing the replication and expression system cherologous gene in a selected host cell. The expression of a heterologous gene sequence operably linked to the selected heterologous gene sequence or the production of selected heterologous sequences and the production of selected betaroins in insect and mammalian host cells e.g. human, rodent and primate cells. EPV thymidine kinase and spheroidin genes can also be used in vertebrate poxviruses such as vaccinia and swinepox virus. The expression vectors can also be used for the control of insect pests through the insertion of a gene encoding an insect toxin into the expression vector which will infect the target pest and produce large quantities of the toxin. Spheroidin and thymidine kinase are nonessential proteins which makes them ideal for the insertion of
                                                                      5
                                                                                                                                            60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel expression system for the expression of heterologous sequences in insect and mammalian host cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Spheroidin; Entomopoxvirus; expression system; replication;
heterologous gene expression; thymidine kinase; poxvirus; vaccinia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Spheroidin protein encoded by Amsacta moorei entomopoxvirus genome.
                                                                      29;
Length 1003;
                                                                      Indels
                                                                                                                                                                                                           441 clkpkvpknlrlwgwildcdt----srfikhmadgsddldldvrln---
                                                                                                                                                                                                                                                                                    111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                             : | | | : | : | : | 483 -rndiclkgaikghytnviileyantypnctlslgnnrfnnvfd 525
                                                              37;
DB 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   swinepox virus; insect pest control; immunity.
                                                                  12; Mismatches
Score 71.5; I
Pred. No. 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Column 61-68; 72pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Moyer RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY30169 standard; Protein; 1003 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (UYFL ) UNIV FLORIDA RES FOUND INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Amsacta moorei entomopoxvirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95US-0544332.
91US-0657584.
92US-0827685.
92WO-US00855.
92US-0991867.
8.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Li Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95US-0544332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hall RL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-457596/38.
Query Match
Best Local Similarity
Matches 26; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAZ10081
```

Search completed: April 9, 2002, 16:56:43 Job time: 477 sec

ò

a



us-09-490-700-41.rsp

```
P10266 homo sapien P31623 sheep pulmo P51517 simian retr P07572 simian maso P07572 simian maso P04026 hamster int P04026 simian retr P0336 mouse intra P0336 mouse intra P1368 mouse intra P03354 rous sarcom P04584 human immun P2407 human immun P19505 simian immu P05962 human immun P05962 human immun P05895 simian immun P1264 equine infe P12451 human immun P1757 human immun P1757 human immun P1787 human immun P1787 human immun P1787 human immun P1787 human immun P1781 human immun P18818 human immun P1881
                                                                                                                        (without alignments)
281.799 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          human immun
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           bovine leuk
chimpanzee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          simian immu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          human t-ce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                      US-09-490-700-41
959
1 FTIPLAEQDCEKFAFTIPAI......SNLFSILRGDSDLNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                April 9, 2002, 17:07:05; Search time 23.68 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P18042
P27973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P03362
P03361
P17283
                                                                                                                                                                                                                                                                                                                                     100059
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                100059 seqs, 36664827 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POL1_HUMAN
POL_JSRV
                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                   OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POL_SRV2
POL_MPMV
POL_IPHA
POL_SRV1
POL_SMRVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POL_RSVP
POL_HV2RO
POL_HV2CA
POL_SIVSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POL_EIAV9
POL_EIAVC
POL_HV2SB
POL_HV2D1
POL_HV2D1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POL_HTL1A
POL_BLVJ
POL_SIVCZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POL_MMTVB
POL_IPMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POL_EIAVY
POL_SIVS4
POL_HV2NZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POL_SIVMK POL_SIVM1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POL_HV2ST
POL_HV2D2
POL_HV2BE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POL_IPMAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POL_BLVAU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POL_SIVA1
                                                                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                  Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  %
Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SwissProt_39:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 895
1036
1036
11022
11022
11022
11023
11035
11035
11035
11058
11058
11058
11058
11058
11058
11058
11058
11058
11058
                                                                                                                                                                         Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score
                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database :
                                                                                                                                                                                                                                                                                                Searched:
                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
```

2 6 7 7 6 6 6 6 6 6 6 6									0;			
immun t-cel t-cel immun immun immun immun immun immun immun immun			63		\$				s	60	120	180
simian immu human immun human immun toel human immun human immun human immun human immun human immun bovine immu human immun			PTASI	nt;	related				Gaps			
			CRI	sto	rel		;	ง ข	0;	PVR III	0 - 0 W - 0	LNS
P27980 P24740 P03363 P14078 P04585 P03368 P20892 P19560 P19561 P04588			REVERSE TRANSCRIPTASE	Euteleostomi ; Homo.	ā			nucrease !;	874; 1s	FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALOPVRDKF 	SDCYIIHYEDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 	RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM
			SE	ng .	ge		3	; ry 64;	Length 87 Indels	TFV	TST	SIL
•			SVER	Craniata; Vertebrata; E Catarrhini; Hominidae;	H.; retrovirus g			ASE.	Len ; I	0 = 10	0K 1Q	SNLF
				ebre	rovi			уше 1РТ <i>I</i> 8В6	1; 81; 2;	NSP1	IASI IASI	YAMS
		A.	e) ate	ert Ho	H.; ret			POL SCR	DB 2e-	GML	GLA	IPT
	SNTS	874	e update) .ion update) [CONTAINS:	a; v ini;	sno ons		;	LIECTEGO DNA POLYME REVERSE TRANSCRIPT ENDONUCLEASE. FD985989798018B6	Score 933; DB 1; Pred. No. 2.2e-81; 3; Mismatches 2;	7.P0	ANA/	21.5
SIVAG HV1U4 HTLV2 HV1L1C HV1DR HV1DR HV1DV BIVOO BIVOO HV1DR HV1DR	ALIGNMENT	~	e ur ton [CC	iate rrhi	hiku genc me.'		3	SE J UCLI 8598	93. No smat	OWK OWK	PAEY 	WIR:
POL_SIVAG POL_HV1U4 POL_HTLV2 POL_HV1H2 POL_HV1H2 POL_HV1PR POL_HV1PV POL_HV1PV POL_HV1PV POL_HV1PV POL_HV1PV POL_HV1PV POL_HV1PV POL_HV1PV	ALI	PRT;	enctat	ran ata	SEQUENCE FROM N.A. MEDLINE-87036922; PubMed=3021993; Ono M., Yasunaga T., Miyata T., Ushikubo Nucleotide sequence of human endogenous the mouse mammary tumor virus genome."; J. Virol. 60:589-598(1986).	c. zn.		VER IDON FD9	ore ed. Mi	TRF	7FL 	NIOI HIOIN
		14	ed) sequanno anno PROT		2199 T., an e	986		. E. R. E.		KEPA KEPA		\$118
			eatest st s	data	=3021 ata T human virus 6).	egra egra seH se.	zn z	6 B	90 90 30 40 31 31	N - N	DKL DKL	DFQ DFQ
1046 1002 982 896 1003 1015 1005 1005 1005 1005		RD;	11. 10, Create 11. 10, Last s 11. 40, Last a TED POL POLYP ENDONUCLEASE]	Homo sapiens (Human), Eukaryota, Metazoa; Chordata; C Mammalla; Eutheria; Primates; C	SEQUENCE FROM N.A. MEDLINE-87036922; PubMed=3 MOD M., Yasunaga T., Miyat Nucleotide sequence of hu the mouse mammary tumor vi the mouse (60:589-598(1986))	Int Int RNA RVT	ase 1. 1	e; otei 0 14 1893	/ Match Focal Similarity 97.3%; Hes 177; Conservative	IPA	ETK 	NJ - E
000,00000000000000000000000000000000000		STANDARD;	10, 10, 2 40,	an).	Puk Fuk nce tum 598(77; 77;	tegr tegr tegr aseH e; 1	767 ;	y rvat	FAFT 	Z = Y	7. L.K
4.4.1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.		ST	el. el. el. ATE	Huma azo eria 6;	N. A 222; ga ' guel guel 89-	010 033 021 004		Poly 6 5 AA	rit	CEK	1100 1100	IRKI IRKI
22222222222			189 (Re 189 (Re 101 (Re IS-RELL	Met Met Suth	369 369 300 300 300 300	1980 1980 1980 1980	552; 552; 552; 075; 078;	58; 58 58 874	nila Co	NEOD NEOD	HYF 	XIE
44 		MAN	P10.265; 01-MAR-1989 01-MAR-1989 20-AUG-2001 RETROVIRUS-R	pier ta; a; E xID	E FF FB Y Cas Yas Tide	20000000000000000000000000000000000000	10000	E Gas	հ Տiո 77;	IPLA IPLA		IKPC
243.5 241.5 241.5 241.5 237.5 237.5 235.5 235.5 235.5 235.5 235.5		L Z E	066; IAR-191 IAR-191 IUG-201 IOVIRUS	sa ryo alia	ENC INE I EO I EO			N N N N N N N N N N N N N N N N N N N	Match ocal ss 17			
		T 1 HUMAN POL1_HUMAN	01-2 01-X 01-X 01-X 00-A	CL.	EDL EDL Nuc	ISSP Inte Inte	fam fam fam fam fam	HAI HAI HAI	y M Lo	128	61 188	121 248
88888888888888888888888888888888888888		RESULT 1 POL1_HUMAN ID POL1_	100 W F O	тти УЗ	- 01 5: O : 12 D	R HSSP, D3346; HNW. R HSSP, P03366; HNW. R InterPro; IPR001037; Integrase_c. R InterPro; IPR003308; Integrase_zn. R InterPro; IPR002156; RNaseH. R InterPro; IPR000477; RVTse.	4 11 11 11 11 11 11 11	- HOOW	Query M Best Lo Matches			
		RES	SEPT TES	80008	RARRER RARRER	222222	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	FT FI	0 2	Qy Db	0y Db	Qy Db

```
ENDONUCLEASE].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |:
LS 307
    1:
306 LS 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LT 182
                                                                                                                      POL_SRV2
P51517;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                  RESULT
                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseeisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                    Sheep pulmonary adenomatosis virus (Jaagsiekte sheep retrovirus)
                                                                                                                                                                                01.JUL-1993 (Rel. 26, Created)
01.JUL-1993 (Rel. 26, Last sequence update)
20.AUG-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN (CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54.7%; Score 525; DB 1; Length 870; 53.3%; Pred. No. 2.1e-42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54; Indels
                                                                                                                                                                                                                                                                                                                                           Viruses; Retroid viruses; Retroviridae; Betaretrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein.
) AA; 99312 MW; CFADEAF8D879C033 CRC64;
                                                                                                                                          870 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR001037; Integrase_C.
Interpro; IPR0010308; Integrase_zn.
Interpro; IPR002156; RNaseH.
Interpro; IPR00156; RNaseH.
Interpro; IPR001584; RNee.
Pfam; PF00552; Integrase; 1.
Pfam; PF00075; Integrase_Zn; 1.
                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M80216; AAA89182.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GNMVJA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00665; rve;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 97; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  870 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P03366; 1HMV
                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF00078;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Endonuclease;
                                                                                                                                                                                                                                                                 ENDONUCLEASE]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; C42740;
HSSP; P03366
181 LT 182
                                       309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 LT 182
                                                                                                                                          POL_JSRV
P31623;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                        11
308 LT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                  POL_JSRV

JD 701-JG

DT 01-JG

DT 01-JG

OS 0-JG

OS 0-JG
                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
Endonuclease; Polyprotein.
SEQUENCE 867 AA; 98038 MW; F64227F7365F3659 CRC64;
                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-87151131; Pubmed-2435057;
Thayer R.M., Power M.D., Bryant M.L., Gardner M.B., Barr P.J.,
Luciw P.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          52.5%; Score 503; DB 1; Length 867; 52.2%; Pred. No. 2.7e-40; live 31; Mismatches 56; Indels
867 AA.
                                                                                                                                                                                                                                                             Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=39068;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Interpro; IPR001037; Integrase_C.
Interpro; IPR003308; Integrase_zn.
Interpro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF02022; Integrase_Zn; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M16605; AAA47562.1; -. HSSP; P04585; 1RTI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000477; RVTse.
InterPro; IPR001584; Rve.
Pfam; PF00552; integrase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 52.5%;
Best Local Similarity 52.2%;
Matches 95; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF00075; rnaseH; 1.
STANDARD;
                                                                                                                                                                                                                                     Simian retrovirus SRV-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00078; rvt;
```

ᠻ

POL_MPMV

```
Simian retrovirus SRV-1.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=11942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interpro; IPR001037; Integrase_C.
Interpro; IPR003308; Integrase_Zn.
Interpro; IPR002156; RNaseH.
Interpro; IPR000477; RVTse.
Interpro; IPR001584; RVse.
Pfam; PF00552; Integrase; 1.
Pfam; PP02022; Integrase_Zn; 1.
                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE-85264989; Pubmed-2991563;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ch 51.4%;
1 Similarity 53.3%;
97; Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00075; rnaseH; 1.
Pfam; PF00665; rve; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00078; rvt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ENDONUCLEASE]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POL_SRV1
P04025:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292 FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POL_SRV1
DDT THE READ OCCOOR OF THE PRESENCE OF THE PRESENCE OCCOOR O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no reatrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cell 45:375-385(1986).
-!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186 KOMYIIHYMDDILIAGKDGQQVLQCFDQLKQELTAAGLHIAPEKVQLQDPYTYLGFELNG
                                                                                                  01-AUG-1988 (Rel. 08, Created)
01-AUG-1988 (Rel. 08, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN (CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     51.7%; Score 496; DB 1; Length 867;
52.7%; Pred. No. 1.2e-39;
ive 27; Mismatches 59; Indels
                                                                                                                                                                                                                                                                                     Simian Mason-Pfizer virus (MPMV).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus
                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. (CLONE 6A).
MEDLINE-86189951; Pubmed-2421920;
Sonigo P., Barker C., Hunter E., Wain-Hobson S.;
"Nucleotide sequence of Mason-Ffizer monkey virus: an immunosuppressive D-type retrovirus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98404 MW; 0F5EDFC77B997935 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  863 AA.
                                                    867 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_zn.
InterPro; IPR00477; RNaseH.
InterPro; IPR000477; RVTse.
InterPro; IPR001584; RNe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00552; integrase; 1.
Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
Pfam; PF00665; rve; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M12349; AAA47711.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Endonuclease; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 52.7%
Matches 96; Conservative
                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; C25839; GNLJMP.
HSSP; P04585; 1RTI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF00078; rvt;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  867 AA;
                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11855;
                                                                                                                                                                                                                          ENDONUCLEASE].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POL_IPHA
P04026;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                 POL_MPMV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     306 LS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POL_IPHA
ID POL_I
AC P0402
```

RESULT

a

ð

ò

ò

ò

```
Ono M., Toh H., Miyata T., Awaya T.;
"Nucleotide sequence of the Syrian hamster intracisternal A-particle
gene: close evolutionary relationship of type A particle gene to
types B and D oncovirus genes.";
J. Virol. 55:387-394(1985).
- MISCELLAMBEDUS: READTHROUGH OF THREE TERMINATORS MAY OCCUR: TAA
BETWEEN CODONS ATT FOR 660-ILE AND AAA FOR 661-LYS, TAG BETWEEN
CODONS TCC FOR 832-SER AND TAT FOR 860-ILE.
CCC FOR 859-PRO AND ATT FOR 860-ILE.
HSSP: P04585: IRTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                 Hamster intracisternal a-particle (IAP-H18).
Viruses; Retroid viruses; Retroviridae; Intracisternal A-particles.
NCBI_TaxID=11752;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33 KNIVPQKIEIRKDHLQTLNDFQKLLGDINWLRPFLKIPSADLKPLFDLLEGEPHISSPRK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-0CT-1986 (Rel. 02, Created)
23-0CT-1986 (Rel. 02, Last sequence update)
20-AGC-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
Endonuclease; Polyprotein.
SEQUENCE 863 AA; 97036 MW; 704AEF7FC6B1D886 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 863;
23-OCT-1986 (Rel. 02, Created)
23-OCT-1986 (Rel. 02, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PUTATIVE POL POLIPPOTEIN [CONTAINS: ENDONUCLEASE; REVERSE TRANSCRIPTASE (EC 2.7.7.49)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Score 493; DB 1;
; Pred. No. 2.4e-39;
29; Mismatches 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       867 AA.
```

```
Endonuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POL_MMTVB
P03365;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POL_MMTVB
ID POL_M
AC P0336
AC 21-JUJ
DT 21-JUJ
DT 20-AU
DE POL P
DE POL P
GN MOUSE
 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the FWBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               246 PKITNOKAVIRKDKLOTLNDFOKLLGDINWLRPYLKLTTADLKPLFDTLKGDSNPNSHRS 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                           deficiency syndrome retrovirus.";
Science 231:1567-1572(1986).
-1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186 KQMYIIHYMDDILIAGKDGQQVLQCFDQLKQELTIAGLHIAPEKIQLQDPYTYLGFELNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-JUL-1986 (Rel. 01, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
20-M02-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
ENDONUCLEASE].
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                               "Nucleotide sequence of SRV-1, a type D simian acquired immune
                                                                                                                                                                                                                                                                                                                                                                        fransferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
           MEDLINE-86151668; PubMed-3306247;
Power M.D., Marx P.A., Bryant M.L., Gardner M.B., Barr P.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                  50.5%; Score 484; DB 1; Length 867; 52.2%; Pred. No. 1.7e-38; ive 27; Mismatches 60; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-89073750; PubMed-3201749;
Oda T., Ikeda S., Watanabe S., Hatsushika M., Akiyama K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POL.
Squirrel monkey retrovirus (SMRV-H) (SMRV-HLB).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus.
                                                                                                                                                                                                                                                                                                                                                                                               98361 MW; ABB06A0B770A6473 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       888 AA.
                                                                                                                                                                                                                                HSSP; P04585; IRTI.
InterPro; IPR01037; Integrase_C.
InterPro; IPR0021308; Integrase_zn.
InterPro; IPR002156; RNaseH.
InterPro; IPR00477; RVTse.
InterPro; IPR001584; Rve.
                                                                                                                                                                                                                                                                                                         Pfam; PF00552; integrase; 1.
Pfam; PF00202; integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
Pfam; PF00665; rve; 1.
                                                                                                                                                                                                          EMBL; M11841; AAA47732.1; -.
                                                                                                                                                                                                                                                                                                                                                                                  Endonuclease; Polyprotein.
SEQUENCE 867 AA; 98361 P
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                        PIR; A03963; GNLJSA.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11856;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 l:
LS 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                            95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POL_SMRVH
P03364;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local
                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
POL_SMRVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          306
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          126
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 SDCYIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   243 QQVFTPRVCLKTDHLKTLNDFQKLLGDIQWLRPYLKLPTSALVPLNNILKGDPNPLSVRA 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                 and gene structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPOGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                    -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : ||:|| |||| | : : | | || : : DEAYILHYMDDILLACDSAEAARACYAHIISCLISYGLKIAPDKVQVSEPFSYLGFELHH
                                                                                                                                                                                                                                 Chiu I.-M., Callahan R., Tronick S.R., Schlom J., Aaronson S.A.; "Major pol gene progenitors in the evolution of oncoviruses."; Science 223:364-370(1984).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-JUL-1986 (Rel. 01, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      fransferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
"Molecular cloning, complete nucleotide sequence, and gene of the provirus genome of a retrovirus produced in a human lymphoblastoid cell line.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99165 MW; B6B2CD09C651B98E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               49.4%; Score 474; DB 1;
49.5%; Pred. No. 1.6e-37;
tive 35; Mismatches 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                899 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mouse mammary tumor virus (strain BR6)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; K01706; AAA46815.1; -.
EMBL, M23385; AAA66453.1; ALT_INIT
PIR; C31827; GNLJHD.
PIR; A05072; A05072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HSSP, P03366; 1HMV.
InterPro; IPR0011037; Integrase_C.
InterPro; IPR002188; Integrase_Zn.
InterPro; IPR002156; RNaseH.
InterPro; IPR000477; RVISe.
                                                                                                                                                                          SEQUENCE OF 595-774 FROM N.A. MEDLINE=84097535; Pubmed=6197754;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF02022; Integrase_2n;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001584; Rve.
Pfam; PF00552; integrase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 49.4%
Best Local Similarity 49.5%
Matches 90; Conservative
                                                                                                         Virology 167:468-476(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00075; rnaseH; 1.
Pfam; PF00665; rve; 1.
Pfam; PF00078; rvt; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   888 AA;
```

S

```
SEQUENCE FROM N.A.
MEDLINE=86286596; PubMed=3016667;
Ymer S., Tucker W.Q.J., Campbell H.D., Young I.G.;
"Nucleotide sequence of the intracisternal A-particle genome inserted 5' to the interleukin-3 gene of the leukemia cell line WEHI-3B.";
Nucleic Acids Res. 14:5901-5918(1986).
-!- MISCELLANEOUS: THIS PARTICLE IS A DEFECTIVE RETROVIRUS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       243 DKIVPQKVEIRRDHLHTTINNFQKLLGDINWLRPFLKIPSAELRPLFWYLEGDPHISSPRT 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mouse intracisternal a-particle (IAP-IL3).
Viruses; Retroid viruses; Retroviridae; Intracisternal A-particles.
NCBL_TaxID=11754;
                                                                                         Mouse intracisternal a-particle (IAP-MIA14).
Viruses; Retroid viruses; Retroviridae; Intracisternal A-particles.
                                                                                                                                                                                                      MEDINE-87311859; PubMed-3041022;
Mietz J.A., Grossman Z., Lueders K.K., Kuff E.L.;
"Nucleotide sequence of a complete mouse intracisternal A-particle
genome: relationship to known aspects of particle assembly and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       183 PSLILLLYMDDILLCHKELTMLQKAYPFLLKTLSQWGLQIATEKVQISDTGQFLGSVVSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          48.1%; Score 461; DB 1; Length 867; 51.1%; Pred. No. 2.7e-36;
ive 28; Mismatches 61; Indels
                                                                                                                                                                                                                                                                                                                     J. Virol. 61:3020-3029(1987).
-!- MISCELLANBOUS: THIS PARTICLE IS A DEFECTIVE RETROVIRUS.
PIR: B26787; GNMSIA.
HSSP: P04585; IKLM.
                         PUTATIVE POL POLYPROTEIN (CONTAINS: ENDONUCLEASE; REVERSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Endonuclease; Polyprotein.
SEQUENCE 867 AA; 97778 MW; 7394B47ED63235B6 CRC64;
  20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-0CT-1989 (Rel. 12, Created)
01-0CT-1989 (Rel. 12, Last sequence update)
01-0V-1995 (Rel. 32, Last annotation update)
PROBABLE POL POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      814 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_zn.
InterPro; IPR00156; RNaseH.
InterPro; IPR000477; RVTse.
InterPro; IPR001584; Rve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00552; integrase; 1.
Pfam; PF00202; integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
Pfam; PF00665; rve; 1.
                                               PRANSCRIPTASE (EC 2.7.7.49)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 51.19
nes 93; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                      NCBI_TaxID=11753;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF00078;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LT 304
                                                                                                                                                                                                                                                                                                    function."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POL_IPMAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POL_IPMAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120 FNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILIVRDKY 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180 QDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQG 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   240 DSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNGDSNPISTRK 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                  -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
                                                                                                                                           mammary
                                                                                                                                                                                                                                                   SEQUENCE OF 578-757 FROM N.A.
MEDLINE-84097535; PubMed-6197754;
Chiu I.-M., Callahan R., Tronick S.R., Schlom J., Aaronson S.A.;
"Major pol gene progenitors in the evolution of oncoviruses.";
Science 223:364-370(1984).
                                                                                                     Moore R., Dixon M., Smith R., Peters G., Dickson C.; Complete nucleotide sequence of a milk-transmitted mouse mammar tumor virus: two frameshlft suppression events are required for translation of gag and pol."; J. Virol. 61:480-490(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00078; rvt; 1.
Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match

48.4%; Score 464; DB 1; Length 899;
Best Local Similarity 50.0%; Pred. No. 1.5e-36;
Matches 91; Conservative 30; Mismatches 61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                625 H -> S (IN REF. 2).
102177 MW; 0FC151C3EBE3C417 CRC64;
Viruses; Retroid viruses; Retroviridae; Betaretrovirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001037; Integrase_C.
InterPro; IPR001308; Integrase_zn.
InterPro; IPR002156; RNaseH.
InterPro; IPR000477; RVTse.
InterPro; IPR001584; Rve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                            MEDLINE-87112944; PubMed-3027377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00552; Integrase; 1.
Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; K01707; AAA46540.1; -. EMBL; M15122; AAA46542.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; C26795; GNMVMM.
HSSP; P03366; 1HMV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     899 AA;
                                                                    SEQUENCE FROM N.A.
                         NCBI_TaxID=11758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Endonuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11
300 LT 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POL_IPMA
P11368;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
```

Pfam;

ð

ð g ö g ð 8

6

RESULT

POL_IPMA ID POL_ AC P113 DT 01-JI

181

and for commercial

Usage by

```
modified and this statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Schwartz D., Tizard R., Gilbert W.;
"Nucleotide sequence of Rous sarcoma virus.";
[Call 32:853-869[1983].
-I- MISCELLANEOUS: AS A RESULT OF BASE VARIATIONS, A DIFFERENT VERSION OF THIS SEQUENCE MAY EXIST HAVING 14-SRP, 116-RRG, 176-RRG, 199-LYS, 247-THR, 304-GL, 474-VAL, 555-GLY, 867-GLY, AND 869-LYS.
-I- MISCELLANEOUS: THIS PROTEIN IS SYMTHESIZED AS A GAG-POL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           84 DCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDTLKTLNDFQK 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24 EPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILCAAETKDKLI 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
20-MUC-2001 (Rel. 40, Last annocation update)
POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
ENDONUCLEASE].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42.3%; Score 406; DB 1; Length 814; 51.6%; Pred. No. 4.6e-31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rous sarcoma virus (strain Prague C).
Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
NCBI_TaxID=11888;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91000 MW; A369620A450F729A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 42.3%; Score 406; DB Best Local Similarity 51.6%; Pred. No. 4.6e Matches 82; Conservative 24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    895
                                                                                                                                                                                                                                                                  HOSPIC TOWNORDS INTEGRASE_C.
InterPro: IPR001308; Integrase_C.
InterPro: IPR001308; Integrase_Zn.
InterPro: IPR001508; RNaseH.
InterPro: IPR0015047; RVTse.
InterPro: IPR001584; Rve.
Pfam; PP00552; Integrase; 1.
Pfam; PP00075; Integrase_Zn; 1.
Pfam; PP00075; Integrase_Zn; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. MEDLINE-83155662; PubMed-6299578;
                                                                                                                                                                                                     EMBL; X04120; CAA27732.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Endonuclease; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                PIR; A23597; GNMSIP.
HSSP; P04585; 1RT1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF00078; rvt;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     814 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POL_RSVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
POL_RSVP
1D LSVP
1D 20-13545
DT 21-JUL
DT 21-JUL
DT 20-AUG
DE POL PO
DE POL POL
DE PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               185
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
```

```
3
(See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X. SAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-184 IN COMPLEX WITH INHIB. MEDLINE-95014470; PubMed=7929352; Chen Z., Li Y., Chen E., Hall D.L., Darke P.L., Culberson C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                173 PSLCMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQREPGVQYLGYKLGS 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 RKIKPOKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POL_HV2RO STANDARD; PRT; 1036 AA.
POL_HV2RO STANDARD;
POL_HV2RO
13-AUG-1987 (Rel. 05, Last sequence update)
13-AUG-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);
REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protease in complex with a reduced amide inhibitor and comparison with HIV-1 protease structures.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tong L., Pav S., Pargellis C., Do F., Lamarre D., Anderson P.C.; "Crystal structure of human immunodeficiency virus (HIV) type 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease; Endonuclease; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human immunodeficiency virus type 2 (isolate ROD) (HIV-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Guyader M., Emerman M., Sonigo P., Clavel F., Montagnier
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98661 MW; 7C28319ED8985465 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'Genome organization and transactivation of the human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Proc. Natl. Acad. Sci. U.S.A. 90:8387-8391(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 86-184.
MEDLINE-93391360; PubMed-8378311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41.1%; Score 394; DB 1;
45.8%; Pred. No. 7.2e-30;
tive 29; Mismatches 66
   entities requires a license agreement (or send an email to license@isb-sib.ch)
                                                                                       EMBL; V01197; -; NOT_ANNOTATED_CDS.
                                                                                                                                               HSSP; P03366; 1HMV.
InterPro; IPR01037; Integrase_C.
InterPro; IPR003308; Integrase_zn.
InterPro; IPR002156; RNaseH.
InterPro; IPR000477; RVTse.
InterPro; IPR001584; RVe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. MEDLINE-87173056; PubMed-3031510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              immunodeficiency virus type 2.";
Nature 326:662-669(1987).
                                                                                                                                                                                                                                                                                                                              Pfam; PF00552; integrase; 1.
Pfam; PF02022; Integrase_Zn; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 45.89
nes 82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                             PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Chen Z., Li Y., Chen E. Shafer J.A., Kuo L.C.;
                                                                                                                   PIR; A03955; GNFV1R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF00078; rvt;
                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00665; rve;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                895 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alizon M
                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
```

```
InterPro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000477; RVTse.
InterPro; IPR001584; Rve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00552; integrase; 1
                                                                                                                                                                                                                                                                                                                                                                67; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
   InterPro; IPR001037;
                                                                                                                                                                                                                                                                              1036 AA;
                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
   DETERMINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POL_HV2CA
                 ACT_SITE
STRAND
                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                         STRAND
                                                   STRAND
                                                                                     STRAND
                                                                                                                       STRAND
                                                                                                                                           STRAND
                                                                                                                                                                                            STRAND
                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                STRAND
                                                                       TURN
                                                                                                                                                                                                                              HELIX
                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                             TURN
                                                                                                         TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 86-184 IN COMPLEX WITH INHIB. MEDLINE-97359919; PubMed-9216835; MedLine-97359919; PubMed-9216835; MedLine-P.L., Wernic D., Abraham A., Anderson P.C., Bogri T., Bousquet Y., Croteau G., Glase I., Lamarre D., Liard F., Paris W., Thibeault D., Pav S., Tong L.; "Potent HIV protease inhibitors containing a novel (hydroxyethyl)amide isostere."; Med. Chem. 40:2164-2176 (1997).
-1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
                                                                                  X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-184 IN COMPLEX WITH INHIB. MEDLINE-95338600; PubMed=7613867; Priestle J.P., Fassler A., Rosel J., Tintelnot-Blomley M., Strop P.,
                                                                                                                                                                                                                                                                                                                                                                                             DETERMINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS50175; ASP_PROT_RETROV; 1.
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
"Crystal structure at 1.9-A resolution of human immunodeficiency virus (HIV) II protease complexed with L-735,524, an orally bloavailable inhibitor of the HIV proteases."; J. Biol. Chem. 269:26344-26348(1994).
                                                                                                                                                       "Comparative analysis of the X-ray structures of HIV-1 and HIV-2 proteases in complex with CGP 53820, a novel pseudosymmetric
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001995; Asp_prot_retrov.
InterPro; IPR001969; Asp_protease.
InterPro; IPR001037; Integrase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                               KNOWN AS THE RETROPEPSIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M15390; -; NOT_ANNOTATED_CDS.
EMBL; X05291; -; NOT_ANNOTATED_CDS.
EMBL; M15390; AAB00764.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR003308; Integrase_zn.
Interpro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PS00141; ASP_PROTEASE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF02022; Integrase_Zn; 1. PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00552; integrase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RVTse.
                                                                                                                                                                                                           Structure 3:381-389(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001584; Rve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21-APR-97.
0; POL$2ROD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-JAN-95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PHV; 15-OCT-94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-APR-92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                03-APR-96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                03-APR-96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-JAN-95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                03-DEC-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-OCT-96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-APR-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-0CT-94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .5-OCT-94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-0CT-94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUL-95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GNLJG2
                                                                                                                                         Grutter M.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          B26262;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                M15390;
                                                                                                                                                                                            Inhibitor."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; A02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUPJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                lidB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4UPJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
HIV;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KARARA BARARA BA
```

```
a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR TGA MAY OCCUR BETWEEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              360 KDVIIIQYMDDILIASDRTDLEHDRVVLQLKELLNGLGFSTPDEKFQKDPPYHWMGYELW 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   300 FSIPLHEDFRPYTAFTLPSVNNAEPGKRYIYKVLPQGWKGSPAIFQHTMRQVLEPFRKAN 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTL-GIPTYAMSNLFSILRGDSDLNSK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA-GLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAR-1992 (Rel. 21, Created)
01-MAR-1992 (Rel. 21, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
POL POL POLYPROTEIN (CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);
REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BELONGS TO PEPTIDASE FAMILY A2; ALSO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29.4%; Score 281.5; DB 1; Length 1036; 37.2%; Pred. No. 4.9e-19; Live 30; Mismatches 78; Indels 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Human immunodeficiency virus type 2 (isolate CAM2) (HIV-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                  117080 MW; 5224E354B1DCC83B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; Retroid viruses; Retroviridae; Lentivirus. NCBI_TaxID=11715;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1034 AA.
                         BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       564-ILE AND 565-GLY.
-!- SIMILARIYY: THE PROTEASE BELONGS TO KNOWN AS THE RETROPEPSIN FAMILY.
PIR: B38475; GNLJCA.
HSSP; P04584; JJLD.
.MEROPS; A02.002; ...
InterPro; IPR001995; Asp_prot_retrov.
InterPro; IPR001995; Asp_prot_retrov.
PROTEASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR003308; Integrase_zn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Integrase_C.
```

œ

```
SEQUENCE OF 1029-1145 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 69; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       177 SK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POL_EIAVY
P03371;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 461 EE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
POL_EIAVY
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                        4
                                                                                                                                                                                                                                                                                                                                                                                                                                                              SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVA-NAGLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            359 SDVIIIQYMDDILLASDRTDLEHDKVVLQLKELLNNLGFSTPDEKFQKDPPYRWMGYELW 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINW---IRPTLGIPTYAMSNLFSILRGDSDLN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=90272009; PubMed=1971917; Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.; "Sequence analysis and acute pathogenicity of molecularly cloned
                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Simian immunodeficiency virus (PBj14/BCL-3 isolate) (sooty mangabey) Viruses; Retroid viruses; Retroviridae; Lentivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POL_SIVSP STANDARD; PRT; 1022 AA.
P19505; Q8B140;
01-FEB-1991 (Rel. 17, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
POL POLYPROTEIN (CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);
REVERSE TRANSCRIFTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AIDS Res. Hum. Retroviruses 8:1179-1187(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92368737; PubMed-1503826; PubMust S., Embretson J.E., Fultz P.N., Mullins J.I.; tve of "Molecular clones from a non-acutely pathogenic derivative of SIVsmmPBj14: characterization and comparison to acutely pathogenic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                   6
                                                                                                             PROSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT_RETROV; 1.
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
NUClease; Transferase; RNA-directed DNA polymerase.
CHAIN 85 183 PROTEASE.
                                                                                                                                                                                                                                                                                                                                    29.2%; Score 280.5; DB 1; Length 1034; 37.9%; Pred. No. 6e-19;
                                                                                                                                                                                                                                                                                                                                                                                   73; Indels
                                                                                                                                                                                                                                      109 109 BY SIMILARITY.
1034 AA; 117195 MW; 3514E566AA6D7C86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                   31; Mismatches
Pfam; PF02022; Integrase_Zn; 1.
                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                   Pfam; PF00075; rnaseH; 1. Pfam; PF00665; rve; 1. Pfam; PF00077; rvp; 1. Pfam; PF00078; rvt; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nature 345:636-640(1990).
                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 69; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11738;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DETERMINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIVSMM-PBj14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EE 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177 SK 178
                                                                                                                                                                                                                                    ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POL_SIVSP

AC DOL SIVSP

DT O1-FE

DT 01-FE

D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AIDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
4
entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             346 PDVTLIQYMDDILIASDRTDLEHDRVVLQLKELLNSIGFSTPEEKFQKDPPFQWMGYELM 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA-GLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINW---IRPTLGIPTYAMSNLFSILRGDSDLN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
20-JUC-1986 (Rel. 04), Last annotation update)
POL POLYPROTEIN (COMTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);
REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-86122873; PubMed-3003905;
Stephens R.M., Casey J.W., Rice N.R.;
"Equine infectious anemia virus gag and pol genes: relatedness to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29.0%; Score 278.5; DB 1; Length 1022; 37.9%; Pred. No. 9.3e-19; artive 30; Mismatches 74; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protease; Endonuclease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Equine infectious anemia virus (isolate Wyoming) (EIAV).
Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBL_TaxID=11672;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96 BY SIMILARITY.
115869 MW; 16DFBEA03F289D6A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA polymerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1145 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF0065; rve; 1.
Pfam; PF00077; rvp; 1.
Pfam; PF00077; rvp; 1.
PROSITE; PS01141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT_RETROY; 1.
AIDS; POlyprotein; Hydrolase; Aspartyl | Nuclease; Transferase; RNA-directed DNA
                        or send an email to license@isb-sib.ch)
                                                                                                                                  HSSP; P04584; 1JLD.
HIV; M31325; POLSSMMPBJ.
MEROPS; A02.002; -.
InterPro; IPR001995; Asp_prot_retrov.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEASE
                                                                                                                                                                                                                                                        InterPro; IPR001969; Asp_protease.
InterPro; IPR00137; Integrase_C.
InterPro; IPR003308; Integrase_zn.
InterPro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam, PF00552; integrase, 1.
Pfam, PF02022; Integrase Zn; 1.
Pfam; PF00075; rnaseH; 1.
                                                                                 EMBL; M31325; AAA47753.1; -.
                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000477; RVTse.
InterPro; IPR001584; Rve.
                                                                                                            EMBL; L03298; AAA4777.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         visna and AIDS virus.";
Science 231:589-594(1986)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96 9
1022 AA;
```

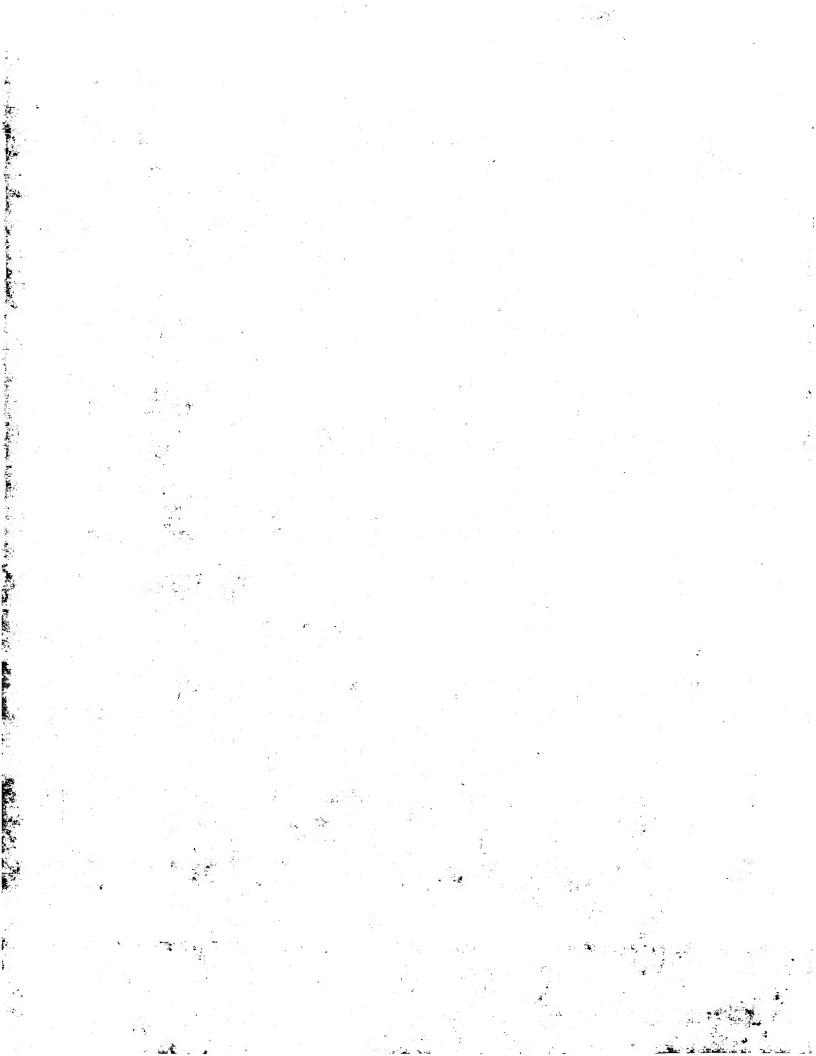
```
RA RUBDILNES-0710553; PubMed-241539;
RA RUBDILNES-0710553; PubMed-241539;
RY RUBDILOW K., Olsen K., Stiegler G., Payne S.L., Montelaro R.C.,
RY Lentivirus genome coganization: the complete nucleotide sequence of
RY Lentivirus genome coganization: the complete nucleotide sequence of
RY Lichtivirus genome coganization: the complete nucleotide sequence of
RY Lichtivirus genome coganization: the complete nucleotide sequence of
RY Lichtivirus genome coganization: the complete nucleotide sequence of
RY Lichtivirus genome coganization: the complete nucleotide sequence of
RY LICHTIVIRUS SECURITION REPORT NUCLES TRUBULES TRUBLIS TR
```

Qy 1 FTIPLAEQDCEKFAPTIPAINNKEPATRFQMKVLPQGMLNSPTICQTFVGRALQPVRDKF 60

11111 : 11111 :

Search completed: April 9, 2002, 17:07:06 Job time: 570 sec

472 ELNOKVIWT 480



	Ltd.
4.5	Compugen
version 4.5	- 2000
GenCore	(c) 1993 -
	Copyright

219241 rotal number of hits satisfying chosen parameters:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries seq length: 0 seq length: 2000000000 Minimum DB Maximum DB

pir1:* pir2:* pir3:* pir4:* PIR_68:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

retrovirus-related pol polyprotein - pol polyprotein (retrovirus-related pol polyprotein - pol polyprotein - pol polyprotein - pol polyprotein - retrovirus related retrovirus related retrovirus retroviral pol polyprotein - pol Description SUMMARIES GNHUER GNNVJA GNLJMP GNLJMP D4 0899 GNLJSA GNLJHD GNNVMM GNNVMM GNNSIA G A48613 S33123 T11560 S30484 GNLJG2 GNLJCA S53092 GNLJEV S30483 GNLJG5 GNLJG3 % Query Match Length DB 1603 843 1019 656 1034 11035 11145 656 11054 11056 11146 11146 Score 933 4493 4494 4 Result Š

GNMVJA

pol polyprotein - sheep pulmonary adenomatosis virus
N;Contains: endonuclease (EC 3.1.-.-); RNA-directed DNA polymerase (EC 2.7.7.49)
S;Species: sheep pulmonary adenomatosis virus
C;Species: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 08-Apr-1994
C;Accession: C42740
R;York, D.F.; Vigne, R.; Verwoerd, D.W.; Querat, G.

RESULT

pol protein - bovi pol polyprotein -	pol polyprotein -	reverse transcript	pol polyprotein -	pol polyprotein -	reverse transcript	pol polyprotein -								
S29358 GNLJST	S08436	GNLJGA	GNLJGG	546347	GNLJGH	GNLJGB	GNLJSI	GNLJH2	GNLJCN	B47175	GNVWLV	GNVWVL	A47175	GNLJBT
7	~	-	_	N	Н	٦	Н	Н	Н	~	Н	Н	~	-
852 1055	1058	852	1035	1039	968	852	1027	985	968	559	1003	1012	559	1053
28.3	28.2	27.9	27.2	26.2	26.0	25.7	25.5	25.1	25.0	24.8	24.8	24.8	24.7	24.6
271 270.5	270.5	268	260.5	251.5	249	246	244.5	241	240	237.5	237.5	237.5	236.5	235.5
30	32	33	34	32	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

GRINGES CHINGTON CARACTERS AND CONTROLLED PROUDGENE - human N.COntains: endonuclease (EC 3.1); RNA directed DNA polymerase (EC 2.7.7.49) C. Species (man) A. Species (

 $^{\rm cl}$

```
246 PKITNQKAVIRKDKLQTLNDFQKLLGDINWLRPYLKLTTGDLKPLFDTLKGDSDPNSHRS 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    51.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 53.38
Matches 97; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Superfamily: pol polyprotein C; Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Cross-references: GB: M73970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-752 <DOR>
                                                                                                                                                                                                                                                                                                                                                                                                                                C; Accession: A03964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: D40899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: A03964
                                                                                                        l:
LS 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Introns: 315/3
                                                                      181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           292 FT 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene: pol
                                                                                                                              306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
            g
                                                                      ò
                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
J. Virol. 66, 4930-4939, 1992
A.Title: Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenous the A.Title: Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenous the A.Accession: C42740
A.Moleoule type: genomic RNA
A.Goment: This protein is likely to be expressed as a gag-pol polyprotein.
C.Comment: The precise boundary between RNA-directed DNA polymerase and endonuclease has C.Goment: pol
A.Goment: pol
A.G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    My Contained and Clone 6A) - Mason-Pfizer monkey virus
pol polyprotein (clone 6A) - Mason-Pfizer monkey virus
pol polyprotein (clone 6A) - Mason-Pfizer monkey virus
C; Species: Mason Pfizer monkey virus
C; Species: Mason Pfizer monkey virus
C; Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C; Accession: C25839
R; Sonigo, P.; Barker, C.; Hunter, E.; Wain-Hobson, S.
Cell 45, 375-385, 1986
A; Title: Nucleotide sequence of Mason-Pfizer monkey virus: an immunosuppressive D-type r
A; Accession: C25839
A; Molecule type: DNA
A; Residues: 1-867 <SON>
A; Molecule type: DNA
A; Residues: 1-867 <SON>
C; Comment: The pol polyprotein contains reverse transcriptase and endonuclease; however, C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A, Gene: pol
C; Superfamily: pol polyprotein
C; Keywords: AIDS; endonuclease; hydrolase; immunodeficiency; nucleotidyltransferase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   246 RVYNTQLVKLQTDHLKTLNDFQKLLGDINWIRPYLKPPYTYLLQPLFDILKGDSDPASPRT 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 870;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 867;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          51.7%; Score 496; DB 1; Length 86 52.7%; Pred. No. 2.8e-38; ive 27; Mismatches 59; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Score 525; DB 1;
; Pred. No. 5.6e-41;
31; Mismatches 54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54.7%;
53.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 52.79 96; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Simi
Matches 97;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |:
LS 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ob
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
```

```
C; Species: Chinese hamster intracisternal A-particle CHIAP34
C; Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 07-Feb-1997
C; Accession: D40899
R; Dorner, A.J.; Bonneville, F.; Kriz, R.; Kelleher, K.; Bean, K.; Kaufman, R.J.
J. Virol, 65, 4713-4719, 1991
A; Tille: Molecular cloning and characterization of a complete Chinese hamster proviru
A; Reference number: A40899; MUID:91333012
                                                                                                                                                                                                                                                                                                                                                                                     A;Molecule type: DNA
A;Residues: 1-863 <000>
C;Comment: Readthrough of three terminators may occur: TAA between codons ATT for 660
59-Pro and ATT for 860-11e.
                                                                                                                                                                                                              R;Ono, M.; Toh, H.; Miyata, T.; Awaya, T.
J. Virol. 55, 387-384, 1985
A;Title: Nucleotide sequence of the Syrian hamster intracisternal A-particle gene:
A;Reference number: A93012; MUID:85264989
retrovirus-related pol polyprotein - golden hamster intracisternal A-particle H18 N;Alternate names: reverse transcriptase C;Species: golden hamster intracisternal A-particle H18 A;Note: host Mesocricetus auratus (golden hamster) C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 30-Sep-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 FIIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pol polyprotein – Chinese hamster intracisternal A-particle CHIAP34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 752;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               51.4%; Score 493; DB 1; L
53.3%; Pred. No. 5.4e-38;
tive 29; Mismatches 56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 490; DB 2;
Pred. No. 8.6e-38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C;Superfamily: pol polyprotein
C;Keywords: polyprotein; reverse transcriptase
```

ö

Gaps

57; Indels

49.4%; Score 474; DB 1; 49.5%; Pred. No. 3.3e-36; tive 35; Mismatches 57;

Length 888;

```
A; Accession: C31827
A; Molecule type: DNA
A; Molecule type: DNA
A; Molecule type: DNA
C; Comment: The pol polyprotein contains reverse transcriptase and endonuclease; howev
C; Cenetics:
A; Genetics:
A; Genetics:
C; Superfamily: pol polyprotein
C; Keywords: endonuclease; hydrolase; nucleotidyltransferase; polyprotein; reverse tra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
123 QQVFTPRVCLKTDHLKTLNDFQKLLGDIQMLRPYLKLPTSALVPLNNILKGDPNPLSVRA 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                          1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
A; Reference number: A31827; MUID:89073750
                                                                                                                                                                                                                                                                                                                                                       90; Conservative
                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ||
LT 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LT 182
                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             =5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        303
                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GNMVMM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Por potyprocein - Similarian Allo Servitus Servitus Structure and Servitus 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pol polyprotein - squirrel monkey retrovirus SMRV-H (SMRV-HLB)

N;Contains: endonuclease (EC 3.1.-..); RNR-directed DNA polymerase (EC 2.7.7.49)

C;Species: squirrel monkey retrovirus SMRV-H

C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 01-May-1998

C;Accession: C31827

R;Oda, T: Ikeda, S:: Watanabe, S:: Hatsushika, M:: Akiyama, K:; Mitsunobu, F.

Virology 167, 468-476, 1988

A;Title: Molecular cloning, complete nucleotide sequence, and gene structure of the prow
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  186 KQMYIIHYMDDILIAGKDGQQVLQCFDQLKQELTIAGLHIAPEKIQLQDPYTYLGFELNG 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RKIKPOKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                  SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                        RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               50.5%; Score 484; DB 1; Length 867, 52.2%; Pred. No. 3.7e-37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
  Indels
:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90;
  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               pol polyprotein - simian AIDS retrovirus SRV-1 N;Contains: endonuclease (EC 3.1....); RNA-dire
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 52.28
Matches 95; Conservative
  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |:
LS 307
97;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LT 182
                                                                                                                                                                                                                                                                                                                                                                                181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                298 FT 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  306
Matches
                                                                                                                                                             61
                                                                                                                                                                                                                                                                     121
                                                                                                     a
                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                             ò
```

```
pol polyprotein - mouse mammary tumor virus
C;Species: mouse mammary tumor virus, MMTV
C;Decies: mouse mammary tumor virus, MMTV
C;Decession: C26795
R;Moore, R.; Dixon, M.; Smith, R.; Peters, G.; Dickson, C.
J. Virol. 61, 480-490, 1987
A;Title: Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus A;Reference number: A39330; MUID:87112944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 899;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    48.4%; Score 464; DB 1; Length 89
50.0%; Pred. No. 2.9e-35;
ive 30; Mismatches 61; Indels
                                                                                                                                                                                                                                                                                                                C;Genetics:
A;Gene: pol
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein; reverse transcriptase
                                                                                                                                                                                                                                    A;Molecule type: DNA
A;Residues: 1-899 <MOO>
A;Cross-references: EMBL:M15122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            l Similarity 50.0%
91; Conservative
```

```
C; Comment: This protein is involved in cell proliferation, differentiation, folliculo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Cross-references: EMBL:M37980; NID:g210272; PIDN:AAA91269.1; PID:g210275
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pol polyprotein - avian leukosis virus
C;Species: avian leukosis virus, ALV
C;Species: avian leukosis virus, ALV
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 26-Aug-1999
C;Accession: S35429, S35435
R;Bieth, E.; Darlix, J.L.
Nucleic Acids Res. 20, 367, 1992
A;Title: Complete nucleotide sequence of a highly infectious avian leukosis virus.
A;Reference number: S35427; MUID:92158628
                                                                                                                                                                                                                                                                                                                                                                                                                           C; Accession: JC7527

K; Graham, K.M.; Ko, C.; Park, K.; Sarge, K.; Park-Sarge, O.K.

B; Gochem. Biophys. Res. Commun. 278, 48-57, 2000

A; Title: Expression of an intracisternal A-particle-like element in rat ovary. A; Reference number: JC7527; MUID: 20525406

A; Contents: Ovary, granulosa cells
A; Accession: JC7527
                                                                                                                                                                                                                                                                                                                            nuclear retroviral polymerase-like protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-2001 #sequence_revision 31-Mar-2001
83 IDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDTLKTLNDFQ 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 ETAYAEVIKTLESNQLFIAPEKVQMGKKGEYLGARITPHNVSPQKIELRKDHLKTLNDFQ 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23 KEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILCAAETKDKL 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A.Status: nucleic acid sequence not shown; translation not shown A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 365;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         143 KLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      42.1%; Score 404; DB 3;
llarity 50.0%; Pred. No. 3.7e-30;
Conservative 24; Mismatches 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 395; DB 2
Pred. No. 8e-29;
8; Mismatches
                                                                                                                                    LLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;Superfamily: pol polyprotein;Keywords: polyprotein; reverse transcriptase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Keywords: ovary; transformation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Cross-references: GB: AA964260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
nes 80; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-895 <BIE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
nes 83; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: mRNA
A; Residues: 1-365 <GRA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Gene: iap-le
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Genetics:
                                                                                                                                                                                                                                                                  1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C; Genetics
                                                                                               144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                              음
                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
          GNMSIA

retrovirus-related pol polyprotein - mouse intracisternal A-particle MIA14

CiSpecies: mouse intracisternal A-particle MIA14

A; Note: host Mus musculus (house mouse)

C; Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 30-Jun-1993

C; Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 30-Jun-1993

C; Accession: B26787

A; Htels: Nucleotide sequence of a complete mouse intracisternal A-particle genome: relat
A; Reference number: A93027; MUD:87311859

A; Accession: B26787

A; Accession: B26787

A; Accession: B26787

A; Mote: the authors translated the codon TGG for residue 64 as Asp, AAC for residue 92 a
C; Comment: The DNA sequence was obtained from GenBank, release 55.0.
C; Comment: This particle is a defective retrovirus.
C; Genetics:
A; Genetics:
C; Seperfamily: pol polyprotein
C; Superfamily: pol polyprotein
C; Keywords: polyprotein; reverse transcriptase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      retrovirus-related pol polyprotein - mouse intracisternal A-particle MIAIL3
C;Species: mouse intracisternal A-particle MIAIL3
A;Note: host Mus musculus (house mouse)
C;Daccession: A23597
C;Accession: A23597
R;Ymer, S: Tucker, W.Q.J.; Campbell, H.D.; Young, I.G.
Nucleic Acids Res. 14, 5901-5918, 1986
A;Title: Nucleotide sequence of the intracisternal A-particle genome inserted, S' to.the A;Reference number: A23597
A;Accession: A23597
A;Accessi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        243 DKIVPQKVEIRRDHLHTHTHTHTPQKLLGDINWLRPFLKIPSAELRPLFWYLEGDPHISSPRT 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84 DCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDTLKTLNDFQK 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EPATREQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILCAAETKDKLI 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 867;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 814;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             53; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Score 461; DB 1;
; Pred. No. 5.3e-35;
28; Mismatches 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Score 406; DB 1;
; Pred. No. 6.7e-30;
24; Mismatches 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Superfamily: pol polyprotein
Keywords: polyprotein; reverse transcriptase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        48.18;
51.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42.3%;
51.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 82; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Simi
Matches 93;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11
LT 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local S:
Matches 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                303
```

a ò a õ g ò QQ

ö

5;

Ω

ŏ

ð

S

us-09-490-700-41.rpr

```
completed: April 9, 2002, 16:58:16 ne: 250 sec
                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 45.3%
Matches 81; Conservative
                                A, Status: preliminary
A, Molecule type: DNA
A, Residues: 1-896 cJOLA
A, Cross-references: GB:L10922
C, Superfamily: pol polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Superfamily: pol polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 39.7%
Best Local Similarity 45.3%
Matches 81; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A Status: preliminary
A Molecule type: DNA
A Residues: 1-806 <100A
A Cross references: GB:L10923
   A; Accession: G48613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search comp
Job time:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E48613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 V.; Optocles: Kous sarcoma virus
A; Note: host Gallus gallus (chicken)
C; Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 16-Jun-2000
C; Accession: A03955; $26418
R; Schwartz, D.; Tizard, R.; Gilbert, W.
Cell 32, 853-869, 1983
A; Title: Nucleotide sequence of Rous sarcoma virus.
A; Reference number: A90834; MUID:81155662
A; Accession: A03955
A; Molecule type: genomic RNA
A; Residues: 1985 scRH>
A; Cross references: GB:VO1197; NID:961695; PID:91335587
A; Residues: GB:VO1197; NID:961695; PID:91335587
A; Residues: GB:VO1197; NID:961695; PID:91335587
A; Residues: as a result of base variations, a different version of this sequence may exist by A; Nobscription: Molecular cloning and DNA sequence analysis of duck-adapted variant of RC A; Reference number: S26417
A; Residues: Lay.
A; Residues: Lay.
A; Recession: S26418
A; Residues: Lay.
A; Recision: S26418
A; Residues: Lay.
A; Residues: Infinity A; Nobscription: Nobscription:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: endonuclease; hydrolase; nucleotidyltransferase; polyprotein; reverse transd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pol polyprotein - myeloblastosis-associated virus (strain MAV-1(N))
(;Species: myeloblastosis-associated virus
(;Species: myeloblastosis-associated virus
(;Accession: 648613
R;Jollot, V.; Boroughs, K.; Lasserre, F.; Crochet, J.; Dambrine, G.; Smith, R.E.; Perbal A;Title: Pathogenic potential of myeloblastosis-associated virus: implication of env pro A;Reference number: A48613; MUID:93331743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pol polyprotein - Rous sarcoma virus
N;Contains: endonuclease (EC 3.1...); RNA-directed DNA polymerase (EC 2.7.7.49)
C;Species: Rous sarcoma virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               173 PSLCMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQREPGVQYLGYKLGS 232
                                                                                                     SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                          233 TYVAPVGL-VAEPRIATLMDVQKLVGSLQWLRPALGIPPRIMGPFYEQLRG-SDPNEAR 289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               233 TYVAPVGL-VAEPRIATLWDVQKLVGSLQWLRPALGIPPRLMGPFYEQLRG-SDPNEAR 289
                                                                                                                                                                                                                                    121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 41.1%; Score 394; DB 1; L
Best Local Similarity 45.8%; Pred. No. 9.9e-29;
Matches 82; Conservative 29; Mismatches 66;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
G48613
                                                                                                  61
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                    ò
                                                                                                     ò
```

```
polyprotein - myeloblastosis-associated virus (strain MAV-2(0)p9)
C;Species: myeloblastosis-associated virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Feb-1998
C;Accession: E48613
R;Joliot, V.: Boroughs, K.; Lasserre, F.; Crochet, J.; Dambrine, G.; Smith, R.E.; Per Virology 195, 812-819, 1993
A;Title: Pathogenic potential of myeloblastosis-associated virus: implication of env A;Reference number: A48613; MuID:93331743
                                                                  3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ä
                                                                                                                                                             SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                  61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 PSLRMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQKEPGVQYLGYKLGS 233
                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                      121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                         1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174 PSLRMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQREPGVQYLGYKLGS
                                                                                                                          1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
         Length 896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 896;
                                                                     92
39.7%; Score 381; DB 2;
45.3%; Pred. No. 1.6e-27;
iive 29; Mismatches 67.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        39.7%; Score 381; DB 2;
45.3%; Pred. No. 1.6e-27;
tive 29; Mismatches 67
```

	•		
	· 4		
A	a · ·		
	<i>A</i> • •		
		52	• 1
		· (c) · . · · · · · · · ·	
			•
		· .	
8			
	**		
(a			
* **			
· .		•	
		· 1	
÷.,			4
	· ·		49
			28
· · · · · · · · · · · · · · · · · · ·		77.	1,4
•			
a ·			0.1
	*		
		*	
• •			
	₹ .	*.*	
	ngerta (n. 1945).		
i i	*		n + 5.00
		÷	
200		×	
w		1). The state of t	
			· ·
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			
		그림을 살아보고 있다.	71
	The second secon		:
·		*	•
	Σ"		
7.	1		
			j
32	· · · · · · · · · · · · · · · · · · ·		
		**************************************	i •
		the state of the s	in the second
			in the state of th

Perfect score:

Sednence:

OM protein

Run on:

Scoring table:

Searched:

Database

Result No.

```
Sequence 4, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 11, Appli
Sequence 1, Appli
                               Sequence 72, Appl
Patent No. 5320958
Sequence 73, Appl
Patent No. 5320958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA-GLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Kraus, Guenter
APPLICANT: Kraus, Guenter
APPLICANT: Wong-Staal, Flossie
APPLICANT: Talbott, Randy
APPLICANT: Poeschla, Eric
TITLE OF INVENTION: Isolation of No. 5883081el HIV-2 Proviruses
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1002 PIKWKLQKIQLPQKEVWTVNDIQKLVGVLNWAAQIYP--GIKT---KHLCKLIRG 1051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINW---IRPTLGIPTYAMSNLFSILRG 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HUMAN IMMUNOBERICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/ISY
HUMAN IMMUNOBERICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/ISY
HUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/331,212
FILING DATE: 03-31-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3080;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28.6%; Score 274.5; DB 6; Length 37.7%; Pred. No. 2.9e-24; Live 31; Mismatches 69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: FRANCHINI, GENOVEFFA; WONG-STAAL, FLOSSIE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: Townsend and Townsend and Crew LLP STREET: Two Embarcadero Center, Eighth Floor CITY: San Francisco STATE: California COUNTRY: USA
                                                                         US-08-679-493A-73
5320958-15
5320958-18
US-09-075-272-4
US-09-603-185-6
US-08-929-967-8
US-08-851-843A-16
US-08-851-843A-16
US-08-851-843A-16
US-08-854-050-16
US-08-854-869-1
US-09-349-546-1
5320958-16
5320958-16
US-08-741-327E-15
US-08-741-32
US-07-743-357-10
US-08-679-493A-72
5320958-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 5, Application US/08659251 Patent No. 5883081 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                   .
65
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity 37.79 S6; Conservative
                                                                                                                                                                                                                                                                                 69
1871
1871
       1003
327
65
237
66
1203
66
379
1079
665
       5223423-4
:Patent No. 5223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO:4:
LENGTH: 3080
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 2
US-08-659-251-5
     209.5
207.5
207.5
141.6
140.5
131.5
1120.5
1119.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.5
114.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GALLO, ROBERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5223423-4
       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 3, Appli
Sequence 71, Appl
Patent No. 5320958
Sequence 14, Appl
Sequence 8, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2, Appli
Patent No. 5320958
Sequence 7, Appli
Sequence 6, Appli
                                                                                                                                                                                                             (without alignments)
117.927 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Patent No. 5223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    App]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sequence 5, Appendix No. 10 May 1 Ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                         959
1 FTIPLAEQDCEKFAFTIPAI......SNLFSILRGDSDLNSKRMLT
                                                                                                                                                                                April 9, 2002, 16:57:29; Search time 34.73 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT US96-11445-5
US-09-238-303-9
US-08-444-842-6
US-08-389-459A-6
US-08-389-459A-6
US-08-117-217-14
US-07-743-357-2
US-07-743-357-4
US-07-743-357-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US95-04910-14
US-07-743-357-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-679-493A-71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-690-265-1
US-07-743-357-9
US-08-463-210-9
US-09-124-900-3
US-07-743-357-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-07-743-357-6
US-07-743-357-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-659-251-5
US-09-256-490-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-07-743-357-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                             212252 segs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5320958-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5320958-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5223423-4
                                                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ü
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                       US-09-490-700-41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       262
913
1005
1016
1016
1016
1016
1004
1016
261
261
260
260
327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
```

4

6

```
61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA-GLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            379 PDVILVQYMDDILIASDRTDLEHDRTVLQLKELLNGLGFSTPDEKFQKDPPYKWMGYELW 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINW---IRPTLGIPTYAMSNLFSILRGDSDLN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match

28.2%; Score 270.5; DB 4; Length 1055;
Best Local Similarity 36.3%; Pred. No. 1.8e-24;
Matches 66; Conservative 32; Mismatches 75; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note= "pol protein encoded by HIV-2KR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: The Regents of the University of California TITLE OF INVENTION: ISOlation of Novel HIV-2 Proviruses NUMBER OF SEQUENCES: 50 CORRESPONDENCE ADDRESS:
                               E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: Robbins, Berliner & Carson
STREET: 201 N. Figueroa Street, 5th Floor
CITY: Los Angeles
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DCCKET NUMBER: 02307E-056410US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
                                                                  COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PATENTIN Release #1.0, CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/256,490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 5, Application PC/TUS9611445 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                          CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/659,251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPOTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OMERATING SYSTEM: PC-DOS/MS-
SOFTWARE: Patentin Release #
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1055 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 1055 amino actor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein
       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ..ocaTION: 1.1055;
CTHER INFORMATION:
US-09-256-490-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA
ZIP: 90012-2628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                          FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        177 SK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EE 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 4
PCT-US96-11445-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINW---IRPTLGIPTYAMSNLFSILRGDSDLN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA-GLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 5, Application US/09256490
| Patent No. 6235881
| GENERAL INFORMATION:
| APPLICANT: Kraus, Guenter | APPLICANT: Wong-Staal, Flossie | APPLICANT: Talbott, Randy | APPLICANT: Toeschila, Eric | APPLICANT: Toeschil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "pol protein encoded by HIV-2KR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32; Mismatches 75; Indels
                                                                                           COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PAtentIn Release #1.0, Version #1.30
SURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,251
FILING DATE: No. 5883081 yet assigned
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28.2%; Score 270.5; DB 2; 36.3%; Pred. No. 1.8e-24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE-DOCKET NUMBER: 30307E-056410US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/001,441
FILING DATE: 26-UUL-1995
ATTORNEY/AGENT INFORMATION:
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1055 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 36.3%
Matches 66; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; LOCATION: 1..1055
; OTHER INFORMATION:
US-08-659-251-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RY: USA
94111-3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      494 EE 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       177 SK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-256-490-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
```

Dp

g ò g

ò

οy

Gaps

. 6

```
356 PEIDLYQYMDDIYIGSDLGKKE---HKQIVEELRKLLLWWGFETPEDKLQEQPPYKWMGY 412
                                                                                                                                                                                                             117 QIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA----GLAIASDKIQTSTPFHYLGM 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 FSDCYIIHYFDDILCAAET----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                       :: || | |: ::::|: || 413 ELYPRKWIQTKELIIPEEPTLNELQKLVGIINWSSQI--IPGLRIKALTNMMKGNQALD
                                   1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24.8%; Score 237.5; DB 1; Length 314; 33.0%; Pred. No. 3.7e-21; Live 32; Mismatches 81; Indels 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Casey D. ENCAPSIDATED POLIOVIRUS NUCLEIC ACACID AND METHODS OF MAKING AND USING SAME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: MOLTOW, CASEY D.
TITLE OF INVENTION: ENCAPSIDATED POLICYTITLE OF INVENTION: ACID AND METHODS OF TITLE OF INVENTION: USING SAME NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS: ADDRESS: ADDRESSE: LAHIVE & COCKFIELD STREET: 60 STATE STREET, SUITE 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,446
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION STAR:
APPLICATION NUMBER: US/08/087,009
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Geary III, William C.
REGISTRATION NUMBER: 31,359
REFERENCE/CDCKET NUMBER: UAG-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 6, Application US/08589446; Patent No. 5614413; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: (617) 227-5941 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         314 amino acids amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 33.0%
Matches 62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE: protein US-08-589-446-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: BOSTON
STATE: MASSACHUSETTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                                                                                                         177 SKRMLT 182
                                                                                                                                                                                                                                                                                                                                                  471 SKRRWT 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
US-08-589-446-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                g
                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                             g
                                                                                                                          ò
                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: Barr, Margaret C.
APPLICANT: Barr, Margaret C.
TITLE OF INVENTION: No. 6284253el Feline Immunodeficiency Virus Nucleotide Sequence FILE REFERENCE: 18617.0059
CURRENT APPLICATION NUMBER: US/09/238,303B
CURRENT FILING DATE: 1999-01-28
EARLIER APPLICATION NUMBER: US 60/072,927
EARLIER FILING DATE: 1998-01-29
NUMBER OF SEQ ID NOS: 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein encoded by the pol gene of a recombinant viral clone constructed from the genomic DNA of a Pallas's cat feline immunodeficiency virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 NRKIKPQKIEIRKDTLKTLNDFQKLLGDINW---IRPTLGIPTYAMSNLFSILRGDSDLN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANA-GLAIASDKIQTSTPFHYLGMQIE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1055;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 24.9%; Score 238.5; DB 4; Length 1150; Best Local Similarity 33.3%; Pred. No. 1.9e-20; Matches 62; Conservative 30; Mismatches 85; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "pol protein encoded by HIV-2KR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 28.2%; Score 270.5; DB 5; Best Local Similarity 36.3%; Pred. No. 1.8e-24; Matches 66; Conservative 32; Mismatches 75;
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/11445
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BEALINER, ROBER: 20,121
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-399C1
TELECOMMUNICATION INFORMATION:
TELEFAX: (213) 977-1001
TELEFAX: (213) 977-1003
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1055 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 9, Application US/09238303B Patent No. 6284253
                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                    ) NAME/KEY: Protein
; LOCATION: 1..1055
; OTHER INFORMATION:
PCT-US96-11445-5
                                                                                                                                                                                                                                                                                                                            amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
US-09-238-303-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Unknown
                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 1150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              177 SK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         494 EE 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-238-303-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

ð g õ ò g ò g Š

```
COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 9
US-08-987-867A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-389-459A-6
                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     153 NPDIVIYQYMDDLYVGSDLEIGQHRTKIEE----LRQHLLRWGLTTPDKKHQKEPPFLWM 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         153 NPDIVIYQYMDDLYVGSDLEIGQHRTKIEE----LRQHLLRWGLTTPDKKHQKEPPFLWM 208
                                  115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                       60 FSDCYIIHYFDDILCAAET-----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 24.8%; Score 237.5; DB 1; Length 314; Best Local Similarity 33.0%; Pred. No. 3.7e-21; Matches 62; Conservative 32; Mismatches 81; Indels 13.
                                                                                                                                                                                                                                                                                                   Casey D.
ENCAPSIDATED POLIOVIRUS NUCLEIC
ALD AND METHODS OF MAKING AND
USING SAME
                                                                                                                                                                                                                     US-08-444-882-6

Sequence 6, Application US/08444882

Patent No. 5622705

GENERAL INFORMATION:
APPLICANT: Morrow, Casey D.
TITLE OF INVENTION: ACID AND METHODS OF
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE 6 COCKFIELD
STREET: 60 STATE STREET, SUITE 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,882
FILING DATE: 19-NAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/087,009
FILING DATE: 01-UUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Geary III, William C.
REFERENCE/DOCKET NUMBER: 31,359
REFERENCE/DOCKET NUMBER: UAG-004
TELECOMMNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: protein US-08-444-882-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: BOSTON
STATE: MASSACHUSETTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           inear
                                                                                                           175 LNSKRMLT 182
                                                                                                                                                267 LTEVIPLT 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: M
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                               g
                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                    ŏ
```

```
60 FSDCYIIHYFDDILCAAET-----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 314;
                                                       Casey D. and Porter, Donna, C. ENCAPSIDATED RECOMBINANT POLIOVIRUS NUCLEIC ACID AND METHODS OF MAKING AND USING SAME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24.8%; Score 237.5; DB 2; 33.0%; Pred. No. 3.7e-21; tive 32; Mismatches 81;
                                                                                                                                                                                                                                                                                                                                   E: LAHIVE & COCKFIELD
60 STATE STREET, SUITE 510
                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/389,459A
FILING DATE: 15-FEB-1995
                                                                                                                                                                                                                                                                          IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
US-08-389-459A-6; Sequence 6, Application US/08389459A; Patent No. 5817512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   314 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 33.0%
Matches 62; Conservative
                                               TITLE OF INVENTION: ENTITLE OF INVENTION: ENTITLE OF INVENTION: USINGBRE OF SEQUENCES: 23 CORRESPONDENCE ADDRESS: ADDRESSEE: LAHIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                                                                                                                                                                                     BOSTON
MASSACHUSETTS
Y: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein
                                                                                                                                                                                                                                                                                          OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                             GENERAL INFORMATION:
APPLICANT: MOLLOW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               267 LTEVIPLT 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE:
                                                                                                                                                                                                                                 02109
```

ŝ

5

```
Sequence 22, Application US/07743357
Patent No. 5858646
GENERAL INFORMATION:
APPLICANT: Kang, Yong C.
TITLE OF INVENTION: Polypeptide having immunological
TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
NUMBERONE OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: KIRBY EADES GALE BAKER
  TITLE OF INVENTION: HIV POSTITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY FITLE OF INVENTION: OF HUMAN HIV STRAINS FILE REFERENCE: 1377-125 CURRENT APPLICATION NUMBER: US/09/117,217 CURRENT FILING DATE: 1998-07-24 NUMBER OF SEQ ID NOS: 15 SOFTWARE: Patentin Ver. 2.0 SEQ ID NO 14 LENGTH: 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60 FSDCYIIHYFDDILCAAET----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                                                                     Length 562;
                                                                                                                                                                                                                                                                                                                                                                                 81; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE: Onterio
COUNTRY: Canada
ZIP: KIM 1H8
COMPUTER REABALE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/743,357
FILING DATE: 21-AUG-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA90/00062
FILING DATE: 23-FEB-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gale, Edwin J.
NAME: Gale, Edwin J.
                                                                                                                                                                                                                                                                                                                                 24.8%; Score 237.5; DB 4; 33.0%; Pred. No. 8.8e-21; iive 32; Mismatches 81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E: KIRBY EADES GALE BAKER
Box 3432, Station D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 237-6900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28,584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: (613) 237-0045
INFORMATION FOR SEQ ID NO: 22:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 33.0%
Matches 62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER:
                                                                                                                                                                                                               ; TYPE: PRT
; ORGANISM: HIV-HXB2
US-09-117-217-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         289 LTEVIPLT 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ottawa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-07-743-357-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET:
                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94 FSVPL-DEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQ 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 FSDCYIIHYFDDILCAAET----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     153 NPDIVIXQYMDDLYVGSDLEIGQHRTKIEE----LRQHLLRWGLTTPDKKHQKEPPFLWM 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 314;
Sequence 6, Application US/08987867A
Patent No. 6053384
GENERAL INFORMATION:
APPLICANT: C. MONTOW et al.
TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
TITLE OF INVENTION: UGLEIC ACID AND METHODS OF MAKING AND
TITLE OF SEQUENCES: 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .7e-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 24.8%; Score 237.5; Best Local Similarity 33.0%; Pred. No. 3.7e Matches 62; Conservative 32; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: UAP-004CPDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEPHONE: (617) 742-4214
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/987,867A
FILING DATE: 09-DEC-1997
CLASSIFICATION 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/087,009
FILING DATE: 01-JUL-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                           MEDLUM TYPE: Floppy disk COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: ASCIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-117-217-14
; Sequence 14, Application US/09117217
; Patent No. 6221578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: de BETHUNE, Marie-Pierre
APPLICANT: HERTOGS, Kurt
APPLICANT: PAUWELS, Rudi
                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 STATE STREET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 314 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                       ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy di
COMPUTER: IBM PC compa
                                                                                                                                                                                                                                     STREET: 28 STATE STRI
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    267 LTEVIPLT 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-987-867A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-07-743-357-4
                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 1, Application US/07743357
Patent No. 585846
ENERAL INFORMATION:
APPLICANT: Kang, Yong C.
TITLE OF INVENTION: Pollypeptide having immunological
TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: KIRBY EABES GALE BAKER
STREET: Box 3432, Station D
                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                     60 FSDCYIIHYFDDILCAAET ----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                    239 NPDIVIYOYMDDLYVGSDLEIGQHRTKIEE----LRQHLLRWGLTTPDKKHQKEPPFLWM 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                   1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                      24.8%; Score 237.5; DB 2; Length 913; 33.0%; Pred. No. 1.8e-20; Live 32; Mismatches 81; Indels 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STATE: COUNTRY: Canada

ZIP: KIM 1H8

COMPUTRY: Canada

ZIP: KIM 1H8

COMPUTRY: Floppy disk

MEDIUM TYPE: Floppy disk

COMPUTRY: TBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NOMBER: US/07/743,357

FLING DATE: 21-AUG-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/CA90/00062

FILING DATE: 33-FBB-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gale, Edwin J.

REGISTRATION NUMBER: 28,584

REGISTRATION NUMBER: 28,584
                                                                                                                                        ORGANISM: Human immunodeficiency virus type 1 STRAIN: HXB2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: 30924-2 TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (613) 237-6900
TELEFAX: (613) 237-0045
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1005 amino acids
                                                                                                                                                                                                                                                                              62; Conservative
                                                     MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
                                                                                                                                                                                                                                        Query Match
Best Local Similarity
amino acid
                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               353 LTEVIPLT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CITY: Ottawa
STATE: Ontari
                  STRANDEDNESS:
                                         TOPOLOGY:
                                                                                                                                                            ; STRAIN: US-07-743-357-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
US-07-743-357-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οχ
```

```
Sequence 4, Application US/07743357
Patent No. 5858646
GENERAL INFORMATION:
TITLE OF INVENTION: Polypeptide having immunological
TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                2
                                                                                                                                                                                                                                                                                                                                                    60 FSDCYIIHYFDDILCAAET ----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                         1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                Indels 13;
                                                                                                                                                                                                                                     Length 1005;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE PATENTING SIZEM: C. DOS. WESTON #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/743,357
FILING DATE: 21-AUG-1991
CLASSIFICATION DATE:
PRIOR APPLICATION DATE:
APPLICATION NUMBER: PCT/CA90/00062
ATONNEY/AGENT INFORMATION:
NAME: Gale, Edwin J.
REGISTRATION NUMBER: 28,584
REGISTRATION NUMBER: 28,584
REGISTRATION NUMBER: 30924-2
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                Query Match 24.8%; Score 237.5; DB 2; Best Local Similarity 33.0%; Pred. No. 2.1e-20; Matches 62; Conservative 32; Mismatches 81;
                                                                                                              ORGANISM: Human immunodeficiency virus type 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: KIRBY EADES GALE BAKER STREET: Box 3432, Station D CITY: Ottawa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1016 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: (613) 237-0045 INFORMATION FOR SEQ ID NO: 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS:
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY: not relevant
.: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE: Ontario
COUNTRY: Canada
ZIP: K1M 1H8
                                    FRAGMENT TYPE: IN ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   445 LTEVIPLT 452
                                                                                                                                       ; STRAIN: HXB2
US-07-743-357-1
                     MOLECULE TYPE:
```

ŝ

```
ORGANISM: Human immunodeficiency virus type 1
STRAIN: BRU
                                                                                                                                                                                                                                                                                                                                                                                          175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                             456 LTEVIPLT 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LTEVIPLT 296
   ORIGINAL SOURCE:
                                        ; STRAIN:
US-07-743-357-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
US-09-690-265-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                               ò
                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
APPLICANT: Kang, Yong C.
TITLE OF INVENTION: Polypeptide having immunological
TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: KIRBY EADES GALE BAKER
STREET: Box 3432, Station D
                                                                                                                                                                           5,
                                                                                                                                                                                                                               60 FSDCYIIHYFDDILCAAET----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                  115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                           1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                            Length 1016;
                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Eloppy disk
COMPUTER: TBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/743,357
FILING DATE: 21-AUG-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA90/00062
FILING DATE: 23-FEB-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gale, Edwin J.
                                                                                                                                       Query Match

24.8%; Score 237.5; DB 2;
Best Local Similarity 33.0%; Pred. No. 2.1e-20;
Matches 62; Conservative 32; Mismatches 81;
   FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Human immunodeficiency virus type 1
STRAIN: PV22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY

REGISTRATION NUMBER: 28,584

REFERENCE/DOCKET NUMBER: 30924-2

TELEPHONE: (613) 237-6900

TELEPHONE: (613) 237-6905

INFORMATION FOR SEQ ID NO: 5: SEQUENCE CHARACTERISTICS:

LENGTH: 1016 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-07-743-357-5; Sequence 5, Application US/07743357; Patent No. 5858646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SS: not relevant not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY: Canada
ZIP: K1M 1H8
HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ontario
                                                                                                                                                                                                                                                                                                                                                                                                                         175 LNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                          456 LTEVIPLT 463
                                                                                    US-07-743-357-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                               음
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
60 FSDCYIIHYFDDILCAAET----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                  342 NPDIVIXQYMDDLYVGSDLEIGQHRTKIEE----LRQHLLRWGLTTPDKKHQKEPPFLWM 397
                                                                                                                                                                                                                                                                                                                                                  115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 FSDCYIIHYFDDILCAAET-----KDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           115 GMQIENRKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSD 174
                                                                                                             1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 FTIPLAEQDCEKF-AFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDK 59
Query Match 24.8%; Score 237.5; DB 2; Length 1016; Best Local Similarity 33.0%; Pred. No. 2.1e-20; Matches 62; Conservative 32; Mismatches 81; Indels 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: UCKUN, Fatih A.
APPLICANT: Wao, Chen
TITLE OF INVENTION: BETA-FLUOROETHYL THIOUREA COMPOUNDS AND USE
FILE REPERENCE: 12152 808C4
CURRENT APPLICATION NUMBER: US/09/690,265
CURRENT FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: US 09/205,167
PRIOR ELILNG DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24.5%; Score 234.5; DB 4; Length 427; 32.4%; Pred. No. 1.4e-20; tive 33; Mismatches 81; Indels 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT : ORGANISM: Human immunodeficiency virus type 1 US-09-690-265-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: April 9, 2002, 16:57:30
Job time: 249 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Sequence 1, Application US/09690265
; Patent No. 6300351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 32.49
Matches 61; Conservative
```

ŝ

MAV reverse transo Full length avian

MAV reverse transc Rous associated vi Peptide #3416 enco Peptide #3509 enco Peptide #3367 enco Sequence of pol pr Sequence encoded b ROD HIV-2 polymera HIV-2 ROD isolate SIVMac339 pol gene SIVMac339 genome p HIV-2 reverse tran Sequence encoded b Sequence encoded b Sequence of clone Deduced sequence e

Sequence of pol pr HIV-2 provirus-enc EIAV pol gene prod PERV-MSN1 BTP-4 re Pol gene product o HIV protease and r Peptide #9570 enco Pol region gene pr HIV-1 pol protein.

Sequence of LAV vi Sequence encoded b Sequence of revers HTLV-III pol prote

Human tPA leader s HIV-1 reverse tran Human immunodefici

ACNPV-HIVYK-pol pr

Perfect score:

Run on:

Seguence:

Scoring table:

Searched:

Database

```
Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human endogenous retrovirus IDDMK1.2-22 pol protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALIGNMENTS
                                                                                       AAM29472
AAM04685
AAP80810
                                                                                                                                              AAP81773
AAR20599
AAX51976
AAR23366
AAW89314
AAB12993
AAP81783
AAP81771
AAP80809
AAR13055
AAB13055
                  AAB12989
AAB12990
AAR80522
AAM16982
                                                                                                                                                                                                                                                                                                                                                                  AAB51187
AAC81068
AAC81068
AAC81068
AAM00180
AAC04791
AAC04793
AAR08053
AAR08053
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW90176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP61507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAW95695 standard; Protein; 182 AA
                                                                                                                                                                13
20
20
20
10
10
                                                                                                                                                                                                                                                                                                                                97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
 8832
8895
88956
6033
10024
10086
10086
10086
11086
11086
11086
11086
11086
11086
11086
11086
11086
                                                                                                                                                                                                                                                                                                                                                                                                                     82
917
313
850
875
902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-097928/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (MACH/) MACH B F.
 diagnosis; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP893691-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-JAN-1999
AAW95695;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW95695
   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human retrovirus-5
Sequence of a poly
Polypeptide with I
MAV reverse transc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Peptide #4562 enco
Peptide #4670 enco
Peptide #4441 enco
                                                                                                                                               (without alignments)
186.335 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human endogenous r
Human endogenous r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Jaagsiekte retrovi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mouse mammary tumo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                   1 FTIPLAEQDCEKFAFTIPAI.....SNLFSILRGDSDLNSKRMLT
                                                                                                                           April 9, 2002, 16:56:45; Search time 72.35 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          **Since **Sinc
                 GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                            fotal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                            522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAW97748
AAM18128
AAM18128
AAM05759
AAY41139
AAY52087
AAF50121
AAF50121
AAF50417
AAF50417
                                                                                         OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAW95695
                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                               US-09-490-700-41
959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A_Geneseq_1101:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182
870
197
197
197
1755
734
775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0
100.0
54.7
50.7
50.7
50.7
50.7
448.4
42.2
42.2
```

959 959 486.5 486.5 464 461 405 397

Score

Result No.

us-09-490-700-41.rag

```
δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                    The sequence is that to an insurint vegetable in the second human endogenous retrovirus (IDDMK1.2-2) probe.

The retrovirus has Superantigen (SAg) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence (without SAg activity but still able to induce an immune response) are useful in vaccines to treat or prevent SAg-related autoimmune similarly to treat such diseases. Retroviral-encoded SAg are important in pathogenesis of autoimmune disease, probably by activating autoreactive T cells. The method is very specific (it can differentiate between expressed and non-expressed viral nucleic acids) and can be used even where the pathogen is an ubiquitous endogenous retrovirus. Blood corporate the pathogen is an ubiquitous endogenous retrovirus. Blood diagnosis can be made before clinical signs are apparent, allowing carly intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HERV; IDDKK1.2-22; superantigen; SAg; antigen; IDDM;
insulin-dependent diabetes mellitus; autoimmune disease; diagnosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .
0
                                                                                      sequence is that of an insulin-dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 182;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human endogenous retrovirus IDDKK1.2-22 polymerase protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0%; Score 959; DB 20; 100.0%; Pred. No. 1.7e-100; tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             therapy; vaccine; polymerase; pol protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW97748 standard; Protein; 182 AA.
preventing autoimmune disease
                                            Claim 31; Fig 7H; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97EP-0401773.
97EP-0112482.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human endogenous retrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 100.
Matches 182; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9905527-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23-JUL-1997;
22-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||
| 1 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-FEB-1999,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW97748;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conrad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW97748
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   THE SECTION OF SECTION
```

```
This is the polymerase protein (Pol) of a new human endogenous retrovirus (HERV), designated IDDMK1.2-22, that has been identified as the source of superantigen (SAG) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAG activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV paving SAG activity (especially including the pre-symptomatic having SAG activity (especially including the pre-symptomatic sepeat, env or pol); (b) a protein or peptide expressed by the HERV capable either (a) the mRNN of an expressed by the HERV capable of blocking transforming specifically associated with the HERV. Products of the invention can be used to identify substances capable of blocking transcription or translation of SAG-encoding nucleic acid sequences, useful in therapy and/or prevention of autoimmune disease associated with the SAG. A nucleic acid encoding the endogenous SAG in IDDM suggests a general model according to which self SAG-driven and systemic activation of autoreactive T cells leads to organ-specific autoimmune disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated human endogenous retrovirus - used to develop products
                                                                                                    for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 182;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JSRV; epithelial carcinoma; ovine; sheep; vaccine; pulmonary adenomatosa; reverse transcriptase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Score 959; DB 20;
llarity 100.0%; Pred. No. 1.7e-100;
Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR31348 standard; Protein; 870 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Jaagsiekte retrovirus Pol protein.
                                                                                                                                                                                                        Claim 9; Fig 7H; 165pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-MAY-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jaagsiekte retrovirus
WPI; 1999-143118/12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
Matches 182; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FR2676455-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAR31348;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
```

3

us-09-490-700-41.rag

```
The present invention relates to human single exon nucleic acid probes (SENP: see AA110068-AA12459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLea cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical espithelial cells. By measuring gene expression, the probes are therefore cervical in grading and/or staging of diseases of the cervix, notably
                                                                                                                                                                                                                                                                                                                                                                                                                                         Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide #4670 encoded by probe for measuring placental gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                           analyzing gene expression in human cervical epithelial cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Probe; microarray; human; placenta; antenatal diagnosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQ 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 486.5; DB 2:
Pred. No. 5.8e-47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 5.86
2; Mismatches
                                                                                                                                                                                                                                                                          Claim 27; SEQ ID No 22954; 487pp; English.
                                                                                                                                                                Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAM30633 standard; Protein; 197
                                                                                                                                (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                Chen W,
2000US-0180312.
2000US-0207456.
2000US-0608408.
2000US-033366.
2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0180312.
2000US-0207456.
2000US-0608408.
2000US-0632366.
2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        50.78;
                                                                                                2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 89.6'
                                                                                                                                                              Hanzel DK,
                                                                                                                                                                                            WPI; 2001-488901/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genetic disorder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200157272-A2.
               26-MAY-2000;
30-JUN-2000;
03-AUG-2000;
21-SEP-2000;
27-SEP-2000;
04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26-MAY-2000;
30-JUN-2000;
03-AUG-2000;
21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-JAN-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAM30633;
                                                                                                                                                              Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
AAM30633
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  246 rvyntqlvklqtdhlktlndfqkllgdinwirpylklptytlqplfdilkgdsdpasprt 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Peptide #4562 encoded by probe for measuring cervical gene expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                  JSRV causes epithelial carcinoma in ovine animals, partic. pulmonary adenomatosa in sheep. The complete cDNA sequence of the JSRV genome was prepared from an approx. 8.7kb band of poly-A RNA isolated from semi-purified lung-lavage samples from infected sheep. The invention includes the Pol amino acid sequence or any part of it which is capable of specific immunological reaction with antibodies directed against JSRV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Probe; human; microarray; gene expression; cervical epithelial cell;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New Jaagslekte Retrovirus and characteristic nucleic acid - also derived proteins, probes and antibodies, useful for in vitro diagnosis and in vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 870;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                          See also AAR31346-7, AAR31349 and AAQ35153-Q35155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 525; DB 14
Pred. No. 2e-50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31; Mismatches
                                                                                            (INRM ) INSERM INST NAT SANTE & RECH MED.
                                                                                                                                                                                                                                                                        Claim 25; Page 39-42; 75pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM18128 standard; Protein; 197 AA
                                                                                                                            Vigne R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54.78;
53.38;
                                91FR-0006060
                                                              91FR-0006060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30-JAN-2001; 2001WO-US00670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 54.7%
Best Local Similarity 53.3%
Matches 97; Conservative
                                                                                                                            Querat GF, Verwoerd D,
                                                                                                                                                           WPI; 1993-020250/03
                                                                                                                                                                                                                                                                                                                                                                                                                                                        870 AA;
                                                                                                                                                                            N-PSDB; AAQ35153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cervical cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200157278-A2.
                              17-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens
                                                              17-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |:
|s 307
20-NOV-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM18128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAM18128
```

셤 à g ò g ò g ä

Gaps

5;

DB 22; Length 197; 4; Indels

```
The invention provides a method for inducing an immune response to mouse mammary tumor virus (MMTV) in a human subject. The method comprises administering to the subject an immunomodulatory composition comprising a pharmaceutical carrier and at least one MMTV antigen (or a discontinuous
                                                                                                                                                                       predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic actiology. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative breast disease and non-carcinoma tumours.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                            (see AA100110-AA110067). The present sequence is a peptide encoded by one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probe hybridises at high stringency to anucleic acid expressed in the human breast. The probes are useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Novel immunomodulatory compositions of mouse mammary tumor virus (MMTV)
                                                                               present invention relates to novel single exon nucleic acid probes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Immune response; mouse mammary tumor virus; MMTV; immunomodulatory; epitope; passive immunotherapy; gag; pol; pro.
                                                                                                                                                                                                                                                                                                                                                                                                                  DB 22; Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mouse mammary tumor virus (MMTV) gag-pol-pro polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         antigens used for treating or preventing MMTV infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQ 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          50.7%; Score 486.5; DB 2; 89.6%; Pred. No. 5.8e-47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Mismatches
                                      Claim 27; SEQ ID No 14499; 322pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY41139 standard; Protein; 1755 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (ITBI-) INT BIOIMMUNE SYSTEMS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig 2; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99WO-US07712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0057003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mouse mammary tumor virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Stewart THM, Gershoni J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-620171/53.
N-PSDB; AAZ23193.
  in a human breast -
                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
nes 95; Conserv
                                                                                                                                                                                                                                                                                                                                                       197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09951268-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14-0CT-1999
                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY41139;
                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                               The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY41139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F X 8 X C C C C C C C C C C C X 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ij
                                                                                                                                                                                                                                                                                          see AAI31315-AAI57546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Probe; human; breast disease; breast cancer; development disorder; inflammatory disease; proliferative breast disease; non-carcinoma tumour.
                                                                                                                                                                                                                                                                           The present invention relates to single exon nucleic acid probes (SENP:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel single exon nucleic acid probe used to measuring gene expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide #4441 encoded by probe for measuring breast gene expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                           Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQ 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 486.5; DB 2:
Pred. No. 5.8e-47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2; Mismatches
                                                                                                                                                                                                                                 Claim 27; SEQ ID No 30902; 654pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DR;
                                                                                                DR;
                                                                                                Rank
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rank
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM05759 standard; Protein; 197
                                                         (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000US-0207456.
2000US-0608408.
2000US-0632366.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        50.7%;
89.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-JAN-2001; 2001WO-US00661.
27-SEP-2000; 2000US-0236359
                2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0234687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                        human genetic disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Penn SG, Hanzel DK,
                                                                                              Penn SG, Hanzel DK,
                                                                                                                                    WPI; 2001-488897/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2001-476286/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
nes 95; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                              197 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200157270-A2
                  04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26-MAY-2000;
30-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         03-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAM05759;
                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

61

ò 셤

셤 ð

RESULT

Matches

ή

Gaps

S.

us-09-490-700-41.rag

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HRV-5; Human retrovirus-5; gag; pro; pol; nucleoprotein; polymerase; recombination; PCR primer; detect; therapy; antibody; vaccine; diagnosis; prognosis; rheumatoid arthritis; osteoarthritis; Sjogren's disease; systemic lupus erythematosus; inflammatory bowel disease; autoimmune disease; inflammatory disease.
epitope of MMTV) in an amount sufficient to elicit an increase in a MMTV specific cellular or humoral response. The compositions and methods of the invention are useful for treating or preventing a human disease or disorder caused by MMTV, or chemical agent that reacts with an antibody directed towards an MMTV antigen. The compositions and antibodies may be used in passive immunotherapy. The immunomodulatory compositions can be used in a preventative manner for those subjects not exposed to MMTV or MMTV antigen reactive agents. The MMTV antigens also have use in diagnostic assays. The present sequence represents the amino acid sequence of the MMTV gag-pol-pro polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human retrovirus-5, its nucleic acid and derived proteins, useful for the treatment, diagnosis and prevention of autoimmune and inflammatory diseases -
                                                                                                                                                                                                                                                                                                             SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                         121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                           1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                              .
0
                                                                                                                                                                                                                           48.4%; Score 464; DB 20; Length 1755; 50.0%; Pred. No. 4.4e-43; tive 30; Mismatches 61; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (CANC-) CANCER RES INST.
(KENN-) KENNEDY INST RHEUMATOLOGY MATHILDA & TER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Boyd MT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human retrovirus-5 pol amino acid sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Venables PJW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY52087 standard; Protein; 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-GB00956.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98GB-0006649.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weiss RA,
                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-601321/51
                                                                                                                                                                              1755 AA;
                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 91; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human retrovirus-5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAZ32523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Griffiths DJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO9950285-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1156 lt 1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-0CT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY52087;
                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       œ
                                                                                                                                                                                                                                                                                                                         916
                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY52087
ID AAY5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8X8888888888X8
                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
This is the human retrovirus-5 HRV-5 pol amino acid sequence. The full length HRV-5 nucleotide sequence AA23523, encodes the gag, pol and progenes of the nucleoprotein of the virus. The pol gane encodes the components of the nucleoprotein of the virus. The pol gane codes for proteins involved in nucleic acid synthesis and recombination, and the progene gives rise to the protease protein. The HRV-5 nucleotide sequences are used in the invention to create PCR primers which can be used to detect HRV-5 in flamed joints, but not in normal synovium. HRV-5 nucleic acids may also be used to screen for specific inhibitors (potential therapeutic agents) and to produce recombinant polypeptides. The virus itself, when disabled, can be used as a gene therapy vector. HRV-5 polypeptides are used to raise antibodies (which may used to detect the virus or as therapeutic inhibitor), to screen for modulators and in vaccines. Fragments of the HRV-5 nucleotide may be used as probes or primers for viral detection (for diagnosis or prognosis) and as sources of the reapeutic antisense sequences. The various therapeutic agents can be resed to treat rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, inflammatory bowel disease, Sjogren's syndrome and other inflammatory or autoimmune conditions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Immunoglobulin E-mediated disease; therapy; B-cell differentiation; immunoglobulin E-binding factor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence of a polypeptide exhibiting mammalian immunoglobulin binding factor activity (IBF) encoded by cDNA clone 23B6p10.2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 734;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                48.1%; Score 461; DB 20; 47.3%; Pred. No. 2.8e-43.ive 37; Mismatches 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAP50121 standard; Protein; 775 AA.
                 Claim 10; Fig 12; 105pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE-) SCHERING BIOTECH CO. (UYJO) JOHNS HOPKINS UNIV. (SCHE) SCHERING-BIOTECH CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85EP-0301834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     84US-0590430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                            734 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-MAR-1984;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-MAR-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27-SEP-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-SEP-1985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      300 lt 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EP155192-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP50121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

Length 775;

9

```
15-JAN-1999;
                                                                                                                                                                                                                                                           30-NOV-2000
                    83;
                                                                                                                                                                                                                                                                                                                                                                                  20-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    procedures
                                                                                                                                                                                                                                     AAB12991;
 Query Match
            Best Local
                      Matches
                                                                                                                                                                                                    AAB12991
                                                                a
                                                                                      ŏ
                                                                                                            q
                                                                                                                                   δλ
                                                                                                                                                     q
                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                             84 DCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDTLKTLNDFQK 143
                                                                                                                                                                                                                                                                                    IBF and IBF cDNA are useful in studies on the immune system. Treatment for IgE-mediated diseases may be possible. IBF may enhance B-cell differentiation into an immunoglobulin-secreting cell.
                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                     EPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILCAAETKDKLI 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cDNA clone coding polypeptide - showing IgE bond factor activity.
                                                                          New complementary DNA clones coding for poly:peptide(s) - with sequence of mammalian immunoglobulin factor and obtd. from transformed or transfected host
                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polypeptide with IgE binding factor activity encoded by clone
                                                                                                                                                                                                                        Length 775;
                                                                                                                                                                                                                                                55; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The polypeptide which shows IgE binding factor activity See also \ensuremath{\mathsf{AAP70416}} .
                                                                                                                                                                                                                                                                                                                                                            144 LLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                         42.2%; Score 405; DB 6; 52.2%; Pred. No. 6.7e-37; tive 21; Mismatches 55;
                      Huff TF;
                      Moore KW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33pp; Japanese.
                                                                                                                     Claim 7; Page 59-64; 71pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP70417 standard; Protein; 775 AA.
(DNAX-) DNAX RES INST MOLEC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85JP-0183810.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85JP-0183810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SCHE ) SCHERING BIOTECH CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                       Query Match
Best Local Similarity 52.2%
Matches 83; Conservative
                      X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Antibody; immunoglobulin
                      Ishizaka
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 29;
                                         WPI; 1985-231863/38.
N-PSDB; AAN50150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1987-097765/14.
                                                                                                                                                                                         775 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAN70646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JP62045600-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-AUG-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-AUG-1985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27-FEB-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP70417;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23B6p102.
                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                               24
g
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                            ò
```

775 AA;

Sequence

```
This invention relates to isolated polynucleotide sequences encoding a polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT) activity. RTs are found in a variety of retroviruses and their defining activity is the ability to synthesise a CDNA strand using an RNA template. The invention includes sequences AAA87808-A87822 and the AAA87842 which represent nucleotide sequences encoding RT polypeptides. Included in the invention are reverse transcriptase protein sequences AAB12989-B12995. Sequences AAA87829-A87839 and AAA87843-A87897 represent oligonucleotides used in the identification and Synthesis of the RT nucleotide sequences of the invention. Which exhibit improved stability and/or improved modified RT proteins, which exhibit improved stability and/or improved polymerase chain reaction (PCR) methods. The nucleotide sequences can be used in sequence are used in improved polymerase chain reaction (PCR) methods. The nucleotide sequence can be used in sequencing methods.
                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Reverse transcriptase; RT; polymerase chain reaction; PCR; retrovirus; stability; solubility.
                                                                                                                                             84 DCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDTLKTLNDFQK 143
                                                                                                                                                                      Gaps
                                                                        24 EPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILCAAETKDKLI 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Novel genes encoding reverse transcriptase polypeptides modified by altering or adding the integrase domains by truncation internally and/or at the C-termini, useful in cDNA synthesis and amplification
                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MAV reverse transcriptase alpha-like polypeptide sequence.
                                    Indels
                                                                                                                                                                                                                                           144 LLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLT 182
42.2%; Score 405; DB 8; 52.2%; Pred. No. 6.7e-37; iive 21; Mismatches 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 2; Page 97-99; 189pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MOLE-) MOLECULAR BIOLOGY RESOURCES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Myeloblastosis-associated virus.
                                                                                                                                                                                                                                                                                                                                                    AAB12991 standard; Protein; 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14-JAN-2000; 2000WO-US00896.
                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-482830/42.
                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     578 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200042199-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Swaminathan N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
```

```
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                             30-NOV-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-JUL-2000.
                                          82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                procedures
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel general altering
                                                                                                                                                                                                                                                                                     AAB12989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   and/or
                                           Matches
                                                                                                                                                                                                                                                 AAB12989
                                                                                                                                                                                                                                                                                     엄
                                                                                                                      ö
                                                                                                                                             g
                                                                                                                                                                       ò
                                                                                                                                                                                              g
                                                                     ò
                              3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This invention relates to isolated polynucleotide sequences encoding a polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT) extivity. RTS are found in a variety of retroviruses and their defining activity is the ability to synthesise a cDNA strand using an RNA template. The invention includes sequences AAA87808 A87822 and AAA8780.487842 which represent nucleotide sequences encoding RT polypeptides. Included in the invention are reverse transcriptase protein sequences AAB12989-B12955. Sequences AAA87829-A87839 and AAA87843-A87897 represent oligonucleotides used in the identification and synthesis of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences encode modified RT proteins, which exhibit improved stability and/or improved solubility, relative to naturally occurring reverse transcriptases. The modified RT nucleotide and protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence represents a reverse transcriptase protein of the invention.
                                                                                                                                                                                                                                                                                                                                                  Reverse transcriptase; RT; polymerase chain reaction; PCR; retrovirus;
                                                                   SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                 173 ps1rmlhymdd111aasshdg1eaageevist1eragftispdkvgrepgvgy1gyk1gs 232
                                                                                                                                                                      1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                       121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genes encoding reverse transcriptase polypeptides modified by ing or adding the integrase domains by truncation internally at the C-termini, useful in cDNA synthesis and amplification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 are used in improved polymerase chain reaction (PCR) methods. The nucleotide sequences can be used in sequencing methods.
                              5;
     Length 578;
                              Indels
                                                                                                                                                                                                                                                                                                                        MAV reverse transcriptase beta-like polypeptide sequence.
   41.4%; Score 397; DB 21;
45.8%; Pred. No. 3.6e-36;
iive 30; Mismatches 65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Example 2; Page 100-103; 189pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MOLE-) MOLECULAR BIOLOGY RESOURCES.
                                                                                                                                                                                                                                                Ą
                                                                                                                                                                                                                                                                                                                                                                                      Myeloblastosis-associated virus.
                                                                                                                                                                                                                                              AAB12992 standard; Protein; 832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-JAN-2000; 2000WO-US00896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0116099
                                                                                                                                                                                                                                                                                                 (first entry)
                              Conservative
                                                                                                                                                                                                                                                                                                                                                               stability; solubility
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-482830/42.
Query Match
Best Local Similarity
Matches 82; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 832 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                WO200042199-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Swaminathan N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15-JAN-1999;
                                                                                                                                                                                                                                                                                               30-NOV-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                         20-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tering or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      and/or at t
procedures
                                                                                                                                                                                                                                                                      AAB12992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Novel
                                                                                                         61
                                                                                                                                                                                                                                                            g
                                                                            g
                                                                                                                                셤
                                                                                                                                                         à
                                                                                                         à
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This invention relates to isolated polynucleotide sequences encoding a polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT) activity. RTs are found in a variety of retroviruses and their defining activity is the ability to synthesise a cDNA strand using an RNA template. The invention includes sequences AAA87808 A87822 and AAA87840.A67842 which represent nucleotide sequences encoding RT polypeptides. Included in the invention are reverse transcriptase protein sequences AAB12989-B12955. Sequences AAA87803 and AAA87843.A87897 represent oligonucleotides used in the identification and synthesis of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences of the invention. Many of the stability and/or improved solubility. relative to naturally occuring reverse transcriptases. The modified RT nucleotide and protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence represents a reverse transcriptase protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Reverse transcriptase; RT; polymerase chain reaction; PCR; retrovirus;
                                                                                                                                                                                                                                                                                                                                    61 SDCYIIHYFDDILCAAETKOKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                       121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genes encoding reverse transcriptase polypeptides modified by ing or adding the integrase domains by truncation internally r at the C-termini, useful in cDNA synthesis and amplification
                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       are used in improved polymerase chain reaction (PCR) methods. The nucleotide sequences can be used in sequencing methods.
                                                                           3
       Length 832;
                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Full length avian reverse transcriptase (RT) protein.
41.4%; Score 397; DB 21;
45.8%; Pred. No. 5.9e-36;
tive 30; Mismatches 65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 1; Page 89-93; 189pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (MOLE-) MOLECULAR BIOLOGY RESOURCES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB12989 standard; Protein; 895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Myeloblastosis-associated virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-JAN-2000; 2000WO-US00896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0116099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             stability; solubility
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-482830/42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200042199-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Swaminathan N;
```

5

Gaps

5;

65; DB 21;

41.4%; Score 397; DB 21; 45.8%; Pred. No. 6.6e-36; tive 30; Mismatches 65

Length 896;

```
1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                             Conservative
                                                                Best Local Similarity
Matches 82; Conserv
            896 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAQ86076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JP07039378-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-JUL-1993;
                                                                                                                                                                                                                                                                                                                                                       09-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10-FEB-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         engineering
                                                   Query Match
                                                                                                                                                                                                                                                                                                                               AAR80522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                                                                            RESULT 15
                                                                                                                                                                                                                                                                                          AAR80522
                                                                                                                                                                                                                                                                                                                    x os
                                                                                                                                                                                  a
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                g
                                                                                                                                                        δý
                                                                                                                                                                                                            à
                                                                3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This invention relates to isolated polynucleotide sequences encoding a polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT) ectivity. RTs are found in a variety of retroviruses and their defining activity is the ability to synthesise a cDNA strand using an RNA template. The invention includes sequences AA87808-A87822 and AAA87840-A87842 which represent nucleotide sequences encoding RT polypeptides. Included in the invention are reverse transcriptase protein sequences AAB12985-B12955. Sequences AAA87829-A87839 and AAA87843-A87897 represent oligonucleotides used in the identification and synthesis of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences not altitude and protein sequence reverse transcriptases. The modified RT nucleotide and protein sequence
                                                                                                                                                                                                                                                                                                                                                                   MAV reverse transcriptase RT protein modified for eukaryotic expression.
                                                                                                                                                                                                                                                                                                                                                                                             transcriptase; RT; polymerase chain reaction; PCR; retrovirus; y; solubility.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       present sequence represents a reverse transcriptase protein of the
                                                                                                       61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                           Gaps
                                                                                        1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                              121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                               Novel genes encoding reverse transcriptase polypeptides modified by altering or adding the integrase domains by truncation internally and/or at the C-termini, useful in cDNA synthesis and amplification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              are used in improved polymerase chain reaction (PCR) methods. The nucleotide sequences can be used in sequencing methods.
                                                                5;
                                    Length 895;
                                                                Indels
                                                                65;
                                    Score 397; DB 21;
Pred. No. 6.6e-36;
                                                              30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 2; Page 93-97; 189pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MOLE-) MOLECULAR BIOLOGY RESOURCES
                                                                                                                                                                                                                                                                                       AAB12990 standard; Protein; 896 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                   Myeloblastosis-associated virus.
                                    41.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-JAN-2000; 2000WO-US00896.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0116099
                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                  Best Local Similarity 45.8
Matches 82; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-482830/42.
895 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200042199-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Swaminathan N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-JAN-1999;
                                                                                                                                                                                                                                                                                                                                           30-NOV-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     procedures
                                                                                                                                                                                                                                                                                                                                                                                                          stability;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Invention.
 Seguence
                                    Query Match
                                                                                                                                                                                                                                                                                                                 AAB12990
                                                                                                                                                                                                                                                                                                                                                                                               Reverse
                                                                                                                                                                                                                                                              RESULT 1
AAB12990
ID AAB1
 S
                                                                                                               Q
                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                   셤
                                                                                                                                                                                              ò
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAQ86076 encodes AAR80522 the Rous associated virus 2 (RAV-2) reverse transcriptase (RT). Using the plasmid pT8RAV the RT could be recombinantly produced in transformed host cells. The RT obtd. could then be used as a reagent for the synthesis of cDNA from mRNA in genetic engineering.
                                                                      61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                          61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKR 179
                                                                                                                                                                                                            1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Reverse transcriptase gene from Rous associated virus · contained in plasmid prBRAV, for synthesis of cDNA rom mRNA in genetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rous associated virus 2; RAV-2; reverse transcriptase; plasmid pTBRAV; cDNA synthesis from mRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rous associated virus 2 (RAV-2) reverse transcriptase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40.4%; Score 387; DB 16;
45.8%; Pred. No. 8.9e-35;
tive 27; Mismatches 68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 2; Pages 6-9; 11pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                       AAR80522 standard; Protein; 895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93JP-0206926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          93JP-0206926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TAKI ) TAKARA SHUZO CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rous associated virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1995-117859/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
nes 82; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        895 AA;
```

```
Ω
```

ò g Search completed: April 9, 2002, 16:56:46 Job time: 480 sec

*

```
RECUENCE FROM N.A.

RECUENCE FROM N.A.

MEDLINE=99445825; PubMed=10516026;

MEDLINE=99445825; PubMed=10516026;

MEDLINE=99445825; PubMed=10516026;

MA Toenjes R.R., Czauderna F., Kurth R.;

"Genome wide screening, cloning, chromosomal assignment and expression of full-length human endogenous retrovirus type K (HERV-K).";

"J. Virol. 73:9187-9195(1999).

"J. Virol. 73:9187-9195(19
015311 homo sapien 015312 homo sapien 015312 homo sapien 015309 homo sapien 015314 homo sapien 071090 human endog 071089 human endog 09473 sheep pulmo 071091 human endog 09473 sheep pulmo 071091 human endog 0947893 human endog 09404 simian retroscope 071292 simian retroscope 071292 simian retroscope 071292 simian retroscope 071292 simian retroscope 071290 simian retroscope 071290 simian retroscope 071291 simian retroscope 071291 simian retroscope 071291 simian retroscope 071291 simian retroscope 071290 simian comman 071213 exogenous m 083393 mouse mamma 091218 exogenous m 083393 mouse mamma 094212 exogenous m 004095 avian leuko 044095 avian leuko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polymerase.
1, 79386 MW; D6728443636546F0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           000007;
01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
POL PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 99.0%; Score 949; DB 12;
Best Local Similarity 98.9%; Pred. No. 1.7e-83;
Matches 180; Conservative 1; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             702 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
VCBI_TaxID=45617;
                       015313
015312
015309
015314
0200811
071009
071009
071009
041292
PR 893
090000
056224
041290
090000
056224
041290
090000
056220
09123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9HBF9
Q04095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q98WW2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                           PRELIMINARY;
    Pfam; PF00078; rvt;
RNA-directed DNA pol
SEQUENCE 702 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00665; rve;
                                                                                                                                                                                                                                                                                                                     551.7
551.7
551.0
551.0
551.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
561.0
    090c07
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
090C07
    QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9qc07 human endog Q9bxr4 homo sapien Q9bxr3 homo sapien O92152 human endog Q9ukh9 homo sapien O92154 human endog Q14273 homo sapien Q9ukh6 homo sapien Q9ukh5 homo sapien Q9ukh5 homo sapien Q9ukh5 homo sapien O92151 human endog Q9ukh5 homo sapien O92151 human endog Q9ukh5 homo sapien O92151 human endog Q9yir5 human endog Q9xjr5 human endog Q6913 human endog Q6913 human endog Q6913 human endog
                                                                                                                                                                                   (without alignments) 375.798 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                        959
1 FTIPLAEQDCEKFAFTIPAI......SNLFSILRGDSDLNSKRMLT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                         April 9, 2002, 17:06:34; Search time 70.84 Seconds
                         GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                            473505 seqs, 146272329 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               092152
092152
092154
092154
092154
014273
091476
09WIK9
09WIK9
09UK93
09
                                                                                                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9QC07
Q9BXR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archea:*
sp_bacteria:*
sp_fung1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                US-09-490-700-41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SPTREMBL_17:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       702
956
956
596
2294
572
740
1361
1755
872
872
956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sb_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.
5.
7.
7.
10.
11.
11.
14.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9949
9941
9941
9937
9933
9932
9929
9929
9911
948
```

Database :

Result

Š

Searched:

Sednence:

Run on:

Gaps

; 0

~

```
596 5
596 AA;
                                                                                                                                     Best Local Similarity
          SEQUENCE FROM N.A.
TISSUE=BLOOD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2001
                                                                                                                                                                                                                                                                                LT 182
                                                                                                                                                                                                                                                                                                    308 LT 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                SEQUENCE
                                                                                                                            Query Match
                                      Meese E.
                                                                                                                                                                                                                                                                                                                                                  092152
092152;
                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                RESULT
  ò
                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                             Pp
                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                  ö
RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                              138 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                          SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 RKIKPOKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        248 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 307
                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      128 FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPQGMLNSPTICQTFVGRALQPVREKF 187
                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                    "Genomic organization of the Human Endogenous Retrovirus HERV-K(HML
                                                                                                                                                                                                                                                                                                                                                                                                                    FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPOGMLNSPTICOTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human).
Vertebrata; Euteleostomi;
Wammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID=9606;
                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                             Reus K., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch N., Meese E.;
                                                                                                                                                                                                                                                                                                                                                                             Score 941; DB 4; Length 956;
Pred. No. 1.5e-82;
1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                  956 AA; 107747 MW; 648619859FC87331 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                               956 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       926
                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                           2.HOM) (ERVK6) on chromosome 7.";
Genomics 72:314-320(2001).
EMBL; AF298587; AAK11553.1;
                                                                                                                                                                                                                                                                                                                                                                              98.1%;
98.4%;
                                                                                                                                                               01-JUN-2001 (TrEMBLrel. 17,
01-JUN-2001 (TrEMBLrel. 17,
01-JUN-2001 (TrEMBLrel. 17,
POLYMERASE (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 98.4
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (TrEMBLrel. 1 (TrEMBLrel. 1 (TrEMBLrel. 1
                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLYMERASE (FRAGMENT).
                                                                                                                                                                                                                (Human)
                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                          NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUN-2001
01-JUN-2001
01-JUN-2001
                                                                                                                                                                                                                Homo sapiens
                                                                                                                                                                                                                                                                        TISSUE-BLOOD
                                                                                        ||
LT 199
                                                                           LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      309
                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5
                                                                                                                                                                                                                                                                                                                                          NON_TER
                                                                                                                                                       Q9BXR4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9BXR3
                                                                                                                                             Q9BXR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9BXR3
                                      121
                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       308
                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
 à
                                      ð
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 RKIKPOKIEIRKDTLKTLNDFOKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 248 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                        by
                         "Genomic organization of the Human Endogenous Retrovirus HERV-K(HML-2.HOM) (ERVK6) on chromosome 7."; Genomics 72:314-320(2001).
EMBL; AF298588; AAK11554.1; -.
                                                                                                                                                                                                                                                                                                                                              1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      188 SDCYIIHYIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            encoded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                              ;
0
Reus K., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 596;
                                                                                                                                                                                                                                            Length 956,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Berkhout B., Jebbink M., Zsiros J.;
"Identification of an active reverse transcriptase enzyme human endogenous HERV-K retrovirus.";
Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
-:- SHMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE TRANSCRIPTASE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                    956 AA; 107847 MW; 146AD510AB0F9071 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59224F02A74F4980 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (TrEMBLrel. 08, Created)
(TrEMBLrel. 08, Last sequence update)
(TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5;
                                                                                                                                                                                                                                         98.1%; Score 941; DB 4; L
98.4%; Pred. No. 1.5e-82;
iive 1; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 937; DB 12
Pred. No. 2e-82;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             596 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            67368 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF080231; AAC63291.1; -.
InterPro; IPR00156; RNaseH.
InterPro; IPR000477; RVTse.
Pfam; PF00075; rnaseH; 1.
Pfam; PF00078; rvt; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97.78;
97.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 97.7
Best Local Similarity 97.8
Matches 178; Conservative
                                                                                                                                                                                                                                                                                         Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1998 (TrEMBLrel. 01-NOV-1998 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLYMERASE (FRAGMENT)
```

```
LT 1232
                                                                                                                                                    LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 178;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                         092154
                                                                                                                                                                                                      ø
                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231
                                                                                                                                                                        1231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291
                                                                                                                                                                                                               092154
                                                  q
                                                                      ò
                                                                                        В
                                                                                                                                qq
                                                                                                                                                    ò
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                õ
                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                          -1- SIMILARITY: BÉLONGS TO ZN-FINGER CCHC TYPE FAMILY.
-1- SIMILARITY: THE PROPEASE BELONGS TO THE PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EUKARYOTIC ASPARTYL PROFEASES FAMILY.

EMBL: AF164611; AAD51793.1; -
                                        'Many human endogenous retrovirus K (HERV-K) proviruses are unique to
RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM
                                                                                                                                                                                                              Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE=99400989; Pubmed=10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT RETROV; 1.
Aspartyl protease; Hydrolase; RNA-directed DNA polymerase; Zinc-finger.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2294 AA; 257099 MW; 425C81FDDF8A80D0 CRC64;
                                                                                                                                                                                                                                                                                                                        -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                         PRT; 2294 AA
                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR001969; Asp_protease.
Interpro; IPR001995; Asp_prot_retrov
Interpro; IPR001428; dUTPase.
                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001428; duTPsase.
InterPro; IPR000332; Gag_plo.
InterPro; IPR000721; Gag_pld.
InterPro; IPR0001037; Integrase_C.
InterPro; IPR001037; Integrase_C.
InterPro; IPR001036; Integrase_C.
InterPro; IPR001584; NaseH.
InterPro; IPR001584; Rve.
InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF00692: dUTPase; 1.
PF01585: G-patch; 1.
PF01587: Gag_D10; 1.
PF0057: gag_D24; 1.
PF00552: integrase; 1.
PF00722: integrase_Zn; 1.
PF00075: rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PD000946; dUTPase; 1.
                                                                                                                                                                                                                                                                                                            Curr. Biol. 9:861-868(1999).
                                                                                                                                                          01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17, GAG-PRO-POL-ENV PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProDom; PD000946; dUTPase; 1
SMART; SM00443; G-patch; 1.
SMART; SM00343; ZnF_C2HC; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00077; rvp; 1.
Pfam; PF00078; rvt; 1.
Pfam; PF00098; zf-CCHC; 2.
                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                   TRANSCRIPTASE
                                                                                                                                                                                                                              NCBI_TaxID-9606;
                                                                                       308
                                                                    LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                =5
                                                                                                                                                                                                                                                                                                      humans.
                                                                                                                                         о9икн9
                                                                                                                                                                                                                                                                                                                                                                                   HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
         187
                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                            121
                                                247
                                                                                       307
                                                                                                                   RESULT
Q9UKH9
                                               g
                                                                    ò
                                                                                     셤
                                                                                                                                           ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ø
                                                                                          1051 FTIPLAEQDCEKFAFTIPAINNKEPATREQWKVLPQGMLNSPTICQTFVGRALQPVREKF 1110
                                                                                                                                                                                  1111 SDCYIIHYIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIEN 1170
                                                                                                                                                                                                                                                                        61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                   121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   transcriptase enzyme encoded by
                                                                     FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 572;
                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Berkhout B., Jebbink M., Zsiros J.;
"Identification of an active reverse transcriptase enzyn
"Identification of an active reverse transcriptase enzyn
human endogenous HERV-K retrovilus.";
Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
TRANSCRIPTASE).
EMBL; ARO80214; AAC63294.1; -.
InterPro; IPR001156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64753 MW; FDD2AAFA37A3A7D0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             092154;
01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-JUN-2010 (TrEMBLrel. 17, Last annotation update)
POLYMERASE (FRACMENT).
  Pred. No. 1.1e-81;
2; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97.3%; Score 933; DB 12;
milarity 97.8%; Pred. No. 4.7e-82;
Conservative 1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 572 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
  97.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RNA-directed DNA polymerase NON_TER 1 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00075; rnaseH; 1.
Pfam; PF00078; rvt; 1.
Best Local Similarity 97.8
Matches 178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         572 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    572
```

RESULT P87890

DB 4; Length 2294;

Score 937;

97.78;

Query Match

```
-!- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
-!- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY ALSO
KNOWN AS THE BUKARYOTIC ASPARTYL PROTEASES FAMILY.
EMBL, AF164613; AAD51796.1; --
InterPro; IPR001995; ASP_prot_act.ov.
InterPro; IPR001995; Asp_prot_retrov.
InterPro; IPR001322; Gag_plo.
InterPro; IPR003322; Gag_plo.
InterPro; IPR000312; Gag_plo.
InterPro; IPR000467; G_patch.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lenz J.; "Many human endogenous retrovirus K (HERV-K) proviruses are unique to
   Nucleotide sequence of human endogenous retrovirus genome related to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90 SDCYIIHYIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIEN 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           150 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSQRI 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
MEDLINE=99400989; PubMed=10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                        Length 1361;
                                                                                                                                                                                                                                                                                                                                                      Query Match 97.3%; Score 933; DB 4; Length 15 Best Local Similarity 97.3%; Pred. No. 1.4e-81; Matches 177; Conservative 3; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                          CEB91B3F407B9498 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CUTT. BIOL. 9:861-868(1999).
-!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
TRANSCRIPTASE).
                                                     -!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE TRANSCRIPTASE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAY-2000 (TrEMBLrel. 13, Created)
1-MAY-2000 (TrEMBLrel. 13, Last sequence update)
1-JUN-2001 (TrEMBLrel. 17, Last annotation update)
GAG-PRO-POL PROFIEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1755 AA
                  the mouse mammary tumor virus genome.";
J. Virol. 60:589-598(1986).
                                                                                    EMBL, M4123; AAA88033.1; -...
InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_Zn.
InterPro; IPR001584; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTse.
Pfam; PP00552; integrase; 1.
                                                                                                                                                                                                                                                                                                      153797 MW;
                                                                                                                                                                                                                Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
Pfam; PF00665; rve; 1.
                                                                                                                                                                                                                                                                                        d DNA polymerase.
1361 AA; 153797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                         Pfam; PF00078; rvt;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                        RNA-directed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 LT 182
                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  210 LT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      о9икн6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        о9икн6
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSQRI 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                            Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97.3%; Score 933; DB 12; Length 740; nilarity 97.3%; Pred. No. 6.5e-82; Conservative 3; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76AD8F9EA35B6BDB CRC64;
                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ono M., Yasunaga T., Miyata T., Ushikubo H.;
 740 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1361 AA.
                                                                                                                  Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID-45617;
                                    Created)
                                                                                                                                                                                                                                                                                                                                                      EMBL, Y10391; CAA71417.1; -.
InterPro; IPR0031308; Integrase_rn.
InterPro; IPR002156; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR00477; RVTSe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-LIVER;
MEDLINE-87036922; Pubmed-3021993;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       740 AA; 83564 MW;
                                   (TrEMBLrel. 03, C
(TrEMBLrel. 03, L
(TrEMBLrel. 17, L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA-directed DNA polymerase.

NON_TER 1 1 1

NON_TER 740 740

SEQUENCE 740 AM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-JUN-2001 (TrEMBLrel. 17, POL/ENV ORF.
 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                 01-MAY-1997 (TrEMBLrel.
01-JUN-2001 (TrEMBLrel.
POL PROTEIN (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00665; rve; 1
Pfam; PF00078; rvt; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 177; Conserv
                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                    01-MAY-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
P87890
P87890;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           014273;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       014273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   œ
```

g

ò g ŏ g

ò

RESULT
014273
10 C014273
10 DT
01 DT
01 DT
01 DT
01 DT
01 CO
02 EU
00 CC
MA
00 CC
MA
01 CO
MA

ö

ö

S

ö

```
InterPro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GAG-POL OR POL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ę
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9WIK9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SO DE RESERVA DE LA COLOR DE L
      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1051 FIIPLAEQDCEKFAFTIPAINNKEPATREQWKVLPQGMLNSPTICQTFVGRALQPVREKF 1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1171 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSQRI 1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 FTIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Toenjes R.R., Czauderna F., Kurth R.; "Full-length human endogenous retrovirus type K (HERV-K) elements encoding Gag, Pol and Env proteins are localised on chromosomes 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97.3%; Score 933; DB 4; Length 1755; 97.3%; Pred. No. 1.9e-81; Live 3; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE (BY SIMILARITY).
-i- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00098; zf-CCHC; 2.
SMART; SM00446; dUTPase; 1.
SMART; SM00443; G-patch: 1.
SMART; SM00343; ZnF_C2HC; 2.
PROSITE: PS00141; ASP_ROYTEASE; 1.
PROSITE: PS50173, ASP_PROYTERTOW; 1.
ASPARTY] Protease; Hydrolase; RNA-directed DNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1755 AA; 195683 MW; DC348F16E456B7BC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSCRIPTASE).
-!- SIMILARITY: TO RETROVIRAL INTEGRASE, C-TERMINUS.
EMBL; Y1784, CAA76885.1; --
InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_Zn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human endogenous retrovirus K.
Viruses: Retrold viruses; Retroviridae.
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
InterPro; IPR003308; Integrase_zn.
InterPro; IPR002156; RNaseH.
                                                                                                                                                 Pfam; PF00692; dUTPase; 1.
Pfam; PF01585; G-patch; 1.
Pfam; PF00607; gag_pl0; 1.
Pfam; PF00607; gag_p24; 1.
Pfam; PF00072; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
                                                                                         InterPro; IPR000477; RVTse.
InterPro; IPR001878; Znf_CCHC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12, 01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 97.35
Best Local Similarity 97.35
Matches 177; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                               PF00078; rvt; 1
                             IPR002156;
IPR001584;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAG-POL OR POL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LT 1232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POL PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zinc-finger
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         O9WJR4
                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DR RELEASE DE LA PROPERTIE DE LA PORTIE DE LA PROPERTIE DE LA PROPERTIE DE LA 
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
Submitted (WAR-1999) to the EMBL/GenBank/DDBJ databases.

Submitted (WAR-1999) to the EMBL/GenBank/DDBJ databases.

-!-PTM: SPECIFIC ENYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS

C:-I-SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE

TRANSCRIPTASE).

C:-I-SIMILARITY: TO RETROVIRAL INTEGRASE, C-TERMINUS.

REMBL: Y17832: CAA76879:1: -.

RICHEPPO: IPR001037: Integrase_C.

RICHEPPO: IPR001154; RV9-.

RICHEPPO: IPR001154; RV9-.

RICHEPPO: IPR001154: RV9-.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Toenjes R.R., Czauderna F., Kurth R.;
"Full-length human endogenous retrovirus type K (HERV-K) elements
encoding Gag, Pol and Env proteins are localised on chromosomes 7 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44 FIIPLAEQDCEKFAFTIPAINNKEPATRFOWKVLPOGMLNSPTICQTFVGRALQPVREKF 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA binding; Endonuclease; Hydrolase; Nucleotidyltransferase;
Polyprotein; RNA-directed DNA polymerase.
SEQUENCE 872 AA; 97876 MW; E5D0A2390060BD5D CRC64;
                                                                                                                                                                                                                                                                           DNA-binding; Endonuclease; Hydrolase; Nucleotidyltransferase;
Polyprotein; RNA-directed DNA polymerase.
SEQUENCE 872 AA; 97873 MW; FAA4267E3B5BBC8B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 872;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97.2%; Score 932; DB 12; Length 8 97.8%; Pred. No. 1e-81; Live 1; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
POL PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      872 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
InterPro; IPR001584; Rve.
InterPro; IPR00047; RYTse.
Pfam; PF00552; integrase; 1.
Pfam; PF00072; Integrase_Zn; 1.
Pfam; PF000075; rnaseH; 1.
Pfam; PF00665; rve; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 97.8
Matches 178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
```

Query Match

Local

Best Loca Matches

=5

104 121

à 셤 δ q òγ Q9UP31

```
-i- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
-i- SIMILARITY: THE PROTEASE BELONGS TO THE PEPTIDASE FAMILY Al; ALSO
KNOWN AS THE ENVARYOTIC ASPARTYL PROTEASES FAMILY.
EMBL, AF164614; AAD51797.1; -.
EMBL, AF164614; AAD51797.1; -.
EINTERPROFILE SAP_PROTEASE.
InterPro; IPR001995; ASP_PROTEASE.
InterPro; IPR001995; ASP_PROTE_FELTOV.
                  SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                     RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                             t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'Many human endogenous retrovirus K (HERV-K) proviruses are unique
FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00443; G-patch; 1.
SMART; SM00343; ZnF_C2HC; 2.
PROSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT_RETROV; 1.
Aspartyl protease; Hydrolase; RNA-directed DNA polymerase; Zinc-finger.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CUTY. Biol. 9:861-868(1999).
-!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
TRANSCRIPTASE).
                                                                                                                                                                                                                                                                                                                                                                  01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
GAG-PRO-POL PROFIEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR003322; Gag_pilo.
InterPro; IPR000451; Gag_pilo.
InterPro; IPR000467; G_palch.
InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_Zn.
InterPro; IPR003156; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR001584; Rve.
InterPro; IPR001878; Znf_CCHC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
MEDLINE=99400989; PubMed=10469592;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00078; rvt; 1.
Pfam; PF00098; zf-CCHC; 2.
ProDom; PD000946; dUTPase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00692; dUTPase; I.
Pfam; PF01585; G-patch; I.
Pfam; PF00237; Gag_p100; I.
Pfam; PF00607; gag_p24; 1.
Pfam; PF00652; integrase; 1.
                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00665; rve; 1
Pfam; PF00077; rvp; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human)
                                                                                                                                                                                                                                ||
LT 309
                                                                                                                                                                                                          LT 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       humans."
                                                                                                                                                                                                                                                                                                               Q9UKH5
ID Q9UKH5
AC Q9UKH5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lenz J.;
                                                                                                                                                                                                                                                                                                13
                                                                                                                                                                        248
                                                                                                                                                                                                          181
                                                                                                                                                                                                                                             308
~
                                128
                                                                   61
                                                                                                                                         121
                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                 g
                                                                   ò
                                                                                                  g
                                                                                                                                     δy
                                                                                                                                                                        Op
                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                      g
                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                    RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'An almost-intact human endogenous retrovirus K on human chromosome
                                                                                  FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa, Chordata; Craniata; Vertebrata, Euteleostomi;
Mammalla; Eutherla; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
Reus K., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch N., Messe E., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch N., Messe E., Sharther characterization of the almost intact human endogenous retrovirus K on human chromosome 7.";
Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
-: SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE TRANSCRIPTASE).
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
              Length 872;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-99178257; Pubmed-10080172;
MAYEr J., Sauter M., Racz A., Scherer D., Mueller-Lantzsch
Meese E.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97.2%; Score 932; DB 4; Length 956;
llarity 97.8%; Pred. No. 1.1e-81;
Conservative 1; Mismatches 3; Indels
                                                 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    956 AA; 107687 MW; E6872987ADCAE41A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
Last sequence update)
Last annotation update)
                DB 12;
              Score 932; DB 12
Pred. No. 1e-81;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                       956 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AF074086; AAD21097.1; -.
EMBL; AF074086; AAF88167.1; -.
InterPro; IPR001037; Integrase_C.
InterPro; IPR003108; Integrase_Zn.
InterPro; IPR002156; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTse.
                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00552; integrase; 1. Pfam; PF02022; Integrase_Zn; 1. Pfam; PF00075; rnaseH; 1. Pfam; PF00665; rve; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, C)
01-MAY-2000 (TrEMBLrel. 13, Li
01-JUN-2001 (TrEMBLrel. 17, Li
01-JUNERASE (FRAGMENT).
POL.
Homo sapiens (Human).
              97.2%;
97.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.";
Nat. Genet. 21:257-258(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RNA-directed DNA polymerase.
NON_TER 1 1
SEQUENCE 956 AA; 107687 MV
                                             178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF00078; rvt; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 178; Conserv
                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                               225
                                                                                                                                                                                                                                                                                              181 LT 182
```

RESULT OCOURNS OF THE PER PROPERTY OCOURNS OCO

Page

us-09-490-700-41.rspt

```
182
                         292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LT 182
=5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   307
                                                                 RESULT
                      q
                                                                                                    AC
DDT
DDT
DDT
DDT
DDT
RRA
RRA
RRA
DDR
DDR
DDR
DDR
FTT
SQ
SQ
SQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                    ø
                                                                                                                                                      1111 SDCYIIHCIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIEN 1170
                                                                                                                                                                                                     1171 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
TISSUE-BONE MARROW;
MEDLINE-99139020; PubMed-9971820;
BETKhout B., Jebbink M., Zsiros J.;
"Identification of an active reverse transcriptase enzyme encoded by in a citient of an active reverse transcriptase enzyme encoded by in a citien of an active reverse transcriptase of J. Virol. 73:2365-2375(1999).
--I SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE TRANSCRIPTASE).
EMBL: AF080232; AAC63292.1; --
InterPro: IPR002156; RNaseH.
InterPro: IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RKIKPOKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                        61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                        121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                     1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF 60
                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
POLYMERASE (FRAGMENT).
Homo sapiens (Human).
Fustaryota: Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                     Length 1879;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  96.9%; Score 929; DB 4; Length 572; 97.3%; Pred. No. 1.2e-81; Live 2; Mismatches 3; Indels
                                                              Indels
F95204AA2E3B10AD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64684 MW; A431D47DCE09C64F CRC64;
                                  Score 932; DB 4; L. Pred. No. 2.6e-81; 1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                   572 AA
208995 MW;
                                 Query Match 97.2%;
Best Local Similarity 97.8%;
Matches 178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RNA-directed DNA polymerase.

NON_TER 1 1

NON_TER 572 572

SEQUENCE 572 AA; 64684 MW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 97.3
Matches 177; Conservative
                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
1879 AA;
                                                                                                                                                                                                                                                                  1231 LT 1232
                                                                                                                                                                                                                                          181 LT 182
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                   090060
                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231
                                                                                                                                                                                                                                                                                                                     090060
S
                                                                                                                                        ò
                                                                                                                                                                셤
                                                                                                                                                                                         ò
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
SEQUENCE FROM N.A.

Berkhout B., Jebbink M., Zsiros J.;

Berkhout B., Jebbink M., Zsiros J.;

Bentification of an active rever transcriptase enzyme encoded by a human endogenous HERVE retrovirus.";

Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.

-! - STMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE EMBL; AF080229; AAC63290.1;

InterPro; IPR002156; RNaseH.

InterPro; IPR004715; RVTSe.

Pfam: PF000775; rnaseH: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 RKIKPQKIEIRKDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 SDCYIIHYFDDILCAAETKDKLIDCYTFLPAEVANAGLAIASDKIQTSTPFHYLGMQIEN 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 FTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          187 SDCYIIHYIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 597;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 96.4%; Score 924; DB 12; Length 5 Best Local Similarity 95.6%; Pred. No. 3.7e-81; Matches 174; Conservative 5; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67446 MW; B5D80527477E0FEC CRC64;
                                                                         Last sequence update)
Last annotation update)
  ¥.
                                                                                                                                   Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9, 2002, 17:06:35
                                                 Created)
                                               08,
08,
                                                                    (TrEMBLrel. 08, (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RNA-directed DNA polymerase.

NON_TER 1 1

NON_TER 597 597

SEQUENCE 597 AA; 67446 MW
PRELIMINARY;
                                    01-NOV-1998 (TrEMBLrel.
01-NOV-1998 (TrEMBLrel.
01-JUN-2001 (TrEMBLrel.
                                                                                                                        POLYMERASE (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search completed: April
Job time: 584 sec
```

ö

			* .		
					A
					3
		**	·		
					**
			1 × 1		
,			· ·		
			•		7.
		*			
			•		
ę	•				
		•			
		$\mathcal{A} = 0$,		
		•			
* .					,
					8
· Comment of the comm	· "	ř.			
	•	*		***	
•		**	y.		•
			v .		
7.					
		•			
		*	,		· · · · · · · · · · · · · · · · · · ·
7	Tag'		9 7 7 8		7 02
1				4	
				'a	
			•	* 1	
		£*		A Company	
00			3 · · · · · · · · · · · · · · · · · · ·	·1	
	of the second se			0	
19 1 V					
* **			••	* =	, of
ar .				· · · · · · · · · · · · · · · · · · ·	
		•			
		9			

the property of the property o

```
Human endogenous r
Human endogenous r
Peptide #486 encod
Peptide #45 encod
Peptide #474 encod
AmEPV Spheroidin p
AmEPV spheroidin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human endogenous r
Human endogenous r
Human endogenous r
                                                                                                                                                                    (without alignments)
185.311 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                  US-09-490-700-40
978
1 WVTPVTWMDNPIEVYVNDSV......PQLLRTNSVVSKCTSESSCR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                        April 9, 2002, 16:56:44; Search time 72.35 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Actions of the control of the contro
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                           522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW97747
AAW95692
AAW95693
AAW97746
AAW97746
AAM26458
AAM01792
AAR29648
                                                                                            OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A_Geneseq_1101:*
1: /STDce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
181
153
153
561
561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1000.0
885.9
885.9
885.9
222.5
222.5
88.2
8.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6:
7:
9:
10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    978
978
840
840
840
220
220
220
220
80.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                             Perfect score:
                                                                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                           Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database
                                                                                                                                        Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ė.
```

80.5 80.5 80.5 80.5 80.5 72 72 72.5 72.5 72.5 72.5 72.5 70.5 70.5 70.5 70.5 70.5 70.5 70.5 70	10.00	22201182221111122211	AAY30169 AAW54347 AAW54347 AAG08013 AAG43910 AAG43909 AAG43908	Spheroidin protes Canine von Willet Canine von Willet
14 80.5 16 79 17 79 18 79 19 79 20 79 22 72 23 72.5 24 72.5 25 72.5 26 72 27 27 28 71 33 70.5 34 70.5 34 70.5 34 70.5 34 70.5 35 70.5 36 69 41 69 42 69 44 69 44 69 44 69 44 69 44 69 44 69 45 69 41 69 41 69 42 69 43 69 44 69 40 69 41 69 41 69 42 69 43 69 44 69 42 69 43 69 44 69 42 69 43 69 44 69 42 69 43 69 44 69 45 69 42 69 43 69 44 69 45 69 47 70.5 33 70.1999 46 69 47 69 48 69 49 69 41 69 41 69 42 69 43 69 44 69 45 69 41 69 41 69 41 69 42 69 43 69 44 69 45 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69 41 69	221111111111111111111111111111111111111	222 222 222 222 222 222 222 222	AAW54347 AAW70557 AAG08013 AAG43910 AAG43909 AAG43909	Canine von Wi Canine von Wi
16 0.99 18 79 19 79 19 79 20 79 22 76 23 72.5 24 72.5 25 72.5 26 71 29 71 31 71.5 29 71 31 70.5 34 70.5 34 70.5 35 70.5 36 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 45 69 46 69 47 70.5 38 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 70.1 39 69.5 39 70.5 39 69 41 69 42 69 43 69 44 69 45 69 44 69 45 69 47 10.5 39 69 41 10.5 39 69 41 69 42 69 43 69 44 69 45 69 47 10.5 39 69 48 50 49 69 49 69 40 6		77777777777777777777777777777777777777	AAG08013 AAG43910 AAG08012 AAG43909 AAG43908	במודוום אסוו אד
17 79 18 79 19 79 20 79 21 72.5 24 72.5 25 72.5 26 71.5 27 71.5 28 71.5 29 71 31 71 32 70.5 33 70.5 34 70.5 35 70.5 36 69.5 36 69.5 36 70 37 70.5 38 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 70.5 39 69.5 39 69.5 39 69.5 39 70.5 39 70.5 30 7		22 22 23 25 25 25 25 25 25 25 25 25 25 25 25 25	AAG43910 AAG08012 AAG43909 AAG43908	Arabidopsis t
18 79 29 79 20 79 21 76 23 72.5 24 72.5 25 72.5 26 71.5 29 71 30 71 31 71 31 71 32 70.5 34 70.5 34 70.5 34 70.5 34 70.5 36 69 40 69 41 69 42 69 44 69 44 69 44 69 45 69 40 69 40 69 41 69 41 69 42 69 43 69 44 69 40 69 41 69 41 69 42 69 43 69 44 69 45 69 40 69 41 69 41 69 42 69 43 69 44 69 45 69 41 69 42 69 43 69 41 69 42 69 43 69 44 69 45 69 40 70		22 22 23 25 21 25 21 25 21	AAG08012 AAG43909 AAG43908	Arabidopsis t
22 79 79 22 76 22 72 5 24 72 5 25 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 72 5 26 20 20 20 20 20 20 20 20 20 20 20 20 20		22 22 23 21 25 21 25 21	AAG43908 AAG43908	Arabidopsis t
21 79 22 76 23 72.5 24 72.5 25 72.5 26 72.5 27 71.5 28 71.5 29 71.5 31 70.5 31 70.5 33 70.5 34 70.5 34 70.5 36 69 40 69 41 69 42 69 43 69 44 69 44 69 45 69 46 69 47 69 48 69 49 69 40 69 41 69 40 69 41 69 42 69 43 69 44 69 45 69 47 69 48 69 49 69 40 69 41 69 40 69 41 69 42 69 43 69 44 69 45 69 47 69 48 69 49 69 41 69 40 69 41 69 41 69 42 69 43 69 44 69 45 69 46 69 47 69 48 69 49 69 49 69 40 69		22 11 22 22 21 22 21	11000011	Arabidopsis t
22 76 23 24 72.5 24 72.5 25 25 72.5 26 72.5 28 72.5 28 72.5 28 72.5 28 72.5 28 72.5 29 71 31 71 32 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 33 70.5 34 70.	® 4 4 4 4 m m m m m m m	25 11 25 27 27 27 27	AAGUDUTT	Arabidopsis t
24 72.5 25 72.5 26 72.5 27 71.5 28 72.5 29 71 31 71 31 71 31 71 32 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 34 70.5 35 70.5 36 69 44 40 69 42 69 42 69 43 69 69 69 69 69 69 69 69 69 69 69 69 69	****	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	AAU14647	Novel bone marrow
25 72.5 26 72.5 27 71.5 28 71.5 29 71.5 31 71.5 31 70.5 34 70.5 34 70.5 35 70.5 36 69 69 40 69 69 41 69 69 42 69 69 44 69 69 44 69 44 69 45 69 46 69 48 69 49 69 40 69 40 69 40 69 41 69 41 69 42 69 43 69 44 69 45 69 46 69 47 69 48 69 49 69 40 69 4	·	25 27 27 27	AAW13826	Yeast transcr
26 72 27 71.5 28 28 29 71 33 71 33 71 33 71 33 70.5 1997; 27 -JAN-1999. 23 -JUL-1997; 23 -JUL-1997; (MACH/) MACH Conrad B, Ma WPI; 1999-097	4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	20 21 21 21	AAR08221	Recombinant a
27 71.3 29 71 29 71 31 71 32 71 33 70.5 34 70.5 35 70.5 36 69.5 37 70.5 38 69.5 39 69.5 39 69.5 39 69.5 39 69.5 39 69.6 41 69.6 42 69.6 43 69.6 44 69.6 44 69.6 44 69.6 44 69.6 44 69.6 44 69.6 44 69.6 45 69.6 46 69.6 47 69.6 48 69.6 49 69.6 49 69.6 40 69.	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	217	AAY41142	Mouse mammary tun
29 71 33 71 33 70.5 34 70.5 34 70.5 34 70.5 35 70.5 35 70.5 35 70.5 36 69 40 69 41 69 69 42 69 44 69 69 44 69 69 44 69 69 44 69 69 44 69 69 44 69 69 45 69 69 69 69 69 6		; ;	AAEU0581 AAG55721	Human protein
30 71 31 71 32 70.5 34 70.5 35 70.5 36 70.5 37 70.5 38 69.5 40 69 41 69 42 69 43 69 44 69 44 69 44 69 44 69 45 69 46 69 46 69 47 69 48 69 48 69 49 69 40 69 41 69 42 69 43 69 44 69 45 69 46 69 47 69 48 69 48 69 49 69 40 70 70 70 70 70 70 70 70 70 70 70 70 70			AAG55720	Arabidopsis t
31 71 32 70.5 34 70.5 35 70.5 36 70.5 37 70.5 38 69.5 40 69 41 69 42 69 43 69 44 69 44 69 44 69 44 69 45 69 44 69 45 69 46 69 47 69 48 69 48 69 49 69 40 69 41 69 42 69 43 69 44 69 44 69 45 69 46 69 47 69 48 69 48 69 49 69 40 70 1999 40 80 1997 40 1997 4		21	AAG55719	Arabidopsis t
33 70.7 34 70.5 35 70.5 36 69.5 38 69.5 38 69.5 39 69.5 40 69 41 69 42 69 43 69 44 69 44 69 44 69 45 69 44 69 45 69 46 69 47 69 48 69 48 69 49 69 40 69 41 1999 42 69 43 69 44 69 45 69 46 69 47 69 48 69		17	AAW03547	ORF-4 protein sec
34 70.5 35 70 36 70 37 70 38 69.5 39 69.5 40 69 42 69 43 69 44 69 45 69 45 69 46 69 47 69 48 69 49 69 49 69 40 69 40 69 40 69 41 69 42 69 43 69 44 69 45 69 45 69 45 69 45 69 45 69 45 69 45 69 46 47 1999 48 69 48 69 49 69 49 69 40 69 4		7 0	AAY49062	Mouse membrane
35 70 36 70 37 70 38 69.5 39 69.5 41 69 42 69 43 69 44 69 44 69 45 69 46 47 69 48 48 48 49 49 49 40 40 40 40 40 40 40 40 40 40 40 40 40		22	AAM39055	Human polypep
36 70 38 40 69 53 44 69 41 69 41 69 44 69 69 44 60 69 69 69 69 69 69 69 69 69 69 69 69 69		21	AAG47392	Arabidopsis t
38 69.7 39 69 40 69 41 69 42 69 43 69 44 69 44 69 45 69 44 69 45 69 46 99 48 69 46 99 48 69 48 69 48 69 49 69 40 69		21	AAG31623	Arabidopsis thal
39 05,3 40 69 41 69 43 69 43 69 44 69 45 69 45 69 44 69 45 69 46 69 47 69 48 69 48 48 48 48 48 48 48 48 48 48 48 48 48		7.	AAG44355	Arabidopsis t
40 69 41 69 42 69 43 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 44 69 69 69 69 69 69 69 69 69 69 69 69 69		7 [AAK88469	reline infect Arabidopsis t
41 69 42 69 43 69 44 69 44 69 45 69 48 69 49 69 49 69 40 69		21	AAG31811	Arabidopsis thal
42 69 44 69 44 69 44 69 45 69 AAW95694 stan AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SAG; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma		21	AAG31810	Arabidopsis t
SULT 1 195694 AAW95694 stan AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SAq; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH COnrad B, Ma		13	AAW74797	+ (
WULT 195694 AAW95694 stan AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SA9; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH COnrad B, Ma		0 0	AAY 60008	
195694 AAW95694 stan AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SAq; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; COnrad B, MaCH COnrad B, Ma		22	AAM39273	Human polypeptide
MAW95694 stan AAW95694 stan AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SA9; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma			ALIGNMENTS	
AAW95694 stan AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SAg; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; COnrad B, Ma				
AAW95694; 08-JUN-1999 Human endogen IDDM; insulin SAq; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; COnrad B, Ma	ard; Protein;	in;	181 AA.	
08-JUN-1999 Human endogen IDDM: insulin SAG; superant diagnosis; en Homo sapiens. EP893691-Al. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma				
08-JUN-1999 Human endogen IDDM; insulin SA9; superant diagnosis; en Homo sapiens. EP893691-A1. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma				
Human endogen IDDM; insulin SAg; superant diagnosis; en Homo sapiens. EP893691-Al. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma	(first entry	ζ,		
IDDM; insulin SAg; superant diagnosis; en Homo sapiens. EP893691-Al. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma	us retrovirus		IDDMK1.2-22 env/fs	(sad) protein.
IDDM; insulin SAg; superant diagnosis; en Homo sapiens. EP893691-Al. 27-JAN-1999. 23-JUL-1997; 23-JUL-1997; (MACH/) MACH Conrad B, Ma WPI; 1999-097			1	
Homo sap EP893691 27-JAN-1 23-JUL-1 23-JUL-1 (MACH/) Conrad I	dependent gen; provi envelope	diab rus;	abetes mellitus; endoges; s; autoimmune disease;	ogenous retrovirus; e; type 1 diabetes;
HOMEO SER EP893691 27-JAN-1 23-JUL-1 23-JUL-1 (MACH/) COnrad I				
27 - JAN - 1 23 - JUL - 1 23 - JUL - 1 23 - JUL - 1 (MACH/) CON rad B				
23-JUL-1 23-JUL-1 (MACH/) CONTAG WPI; 199				
23-JUL-1 23-JUL-1 (MACH/) COnrad H				
23-JUL-1 23-JUL-1 (MACH/) Conrad E				
23-JUL-] (MACH/) Conrad B	97EP-040177	773.		
(MACH/) Conrad F	97EP-040177	773.		
Conrad F				
WPI; 199	h B;			
WPI; 199				
N-PSDB;	28/09. 16.			
AX Pr Diagnosing human	an antoimmine	9011	disease by detecting	ng retrovirus with
				d with type 1
	ч		~	transf

```
AAW95692
       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ΩQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DXXXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οý
                                                                                                  The Saydance of all tributions of a saydance of a saydance
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSVVSKCTSESSC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 mvtpvtwmdnpievyvndsvwvpgptddrcpakpeeegmminisigyhyppiclgrapgc 60
proteins, antibodies and specific binding agents, used for treating
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HERV; IDDKKI.2-22; superantigen; SAg; antigen; IDDM; insulin-dependent diabetes mellitus; autoimmune disease; diagnosis; therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MVTPVTWMDNPIEVXVNDSVWVPGPTDDRCPAKPEEGGMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                         an insulin-dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human endogenous retrovirus IDDKK1.2-22 Env/F-S (SAg)
                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 978; DB 20;
100.0%; Pred. No. 6.7e-106;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17..19 /note= "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW97747 standard; Protein; 181 AA.
                    preventing autoimmune disease
                                                      Claim 31; Fig 7G; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human endogenous retrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 100.
Matches 181; Conservative
                                                                                         sequence is that of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42..45
                                                                                                                                                                                                                                                                                                                                                                                        181 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO9905527-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAY-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-FEB-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 R 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 r 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW97747;
                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW97747
   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
This is the envelope (Env)/FS polypeptide of a new human endogenous retrovirus (HERN), designated IDDMK1.2-22, that has been identified as the source of superantigen (SAg) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAg activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV having SAg activity comprises specifically detecting in a biological sample either: (a) the mERN of a expressed HERV having SAg activity (especially IDDMK1.2-22 5'LTR, 3' long terminal crepeat, env or pol); (b) a protein or peptide expressed by the HERV (see AAW97745-48); (c) antibodies specifically associated with the HERV. HERV, or (d) SAg activity specifically associated with the HERV. HERV, or (d) SAg activity specifically associated with the HERV. HERV, or (d) SAg activity specifically associated with the HERV. HERV. Or (d) SAG activity specifically associated with the HERV. HERV. Or (d) SAG activity or or of dentify substances here.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSVVSKCTSESSC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          capable of blocking transcription or translation of SAg-encoding nucleic acid sequences, useful in therapy and/or prevention of autoimmune disease associated with the SAg. A nucleic acid encoding human retroviral SAg can be used as a DNA vaccine. Expression of the endogenous SAg in IDDM suggests a general model according to which self SAg-driven and systemic activation of autoreactive T cells leads to organ-specific autoimmune disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated human endogenous retrovirus – used to develop product: for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human endogenous retrovirus IDDMK1.2-22 env protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100.0%; Score 978; DB 20;
100.0%; Pred. No. 6.7e-106;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW95692 standard; Protein; 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 6; Fig 7G; 165pp; English.
97EP-0401773.
97EP-0112482.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                     WPI; 1999-143118/12.
N-PSDB; AAX07191.
                                                                                                                   (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                              B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AA;
23-JUL-1997;
22-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 R 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 <sup>†</sup> 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAW95692
                                                                                                                                                                                              Conrad
```

ö

 \sim

33-JUL-1997;

Homo sapiens EP893691-A1 27-JAN-1999

```
1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                HERV; IDDKK1.2-22; superantigen; SAg; antigen; IDDM; insulin-dependent diabetes mellitus; autoimmune disease; diagnosis; therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New isolated human endogenous retrovirus - used to develop product for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        human retroviral SAg can be used as a DNA vaccine. Expression of the endogenous SAg in IDDM suggests a general model according to which self SAg-driven and systemic activation of autoreactive T cells leads to organ-specific autoimmune disease.
                                  Human endogenous retrovirus IDDKK1.2-22 envelope protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85.9%; Score 840; DB 20;
100.0%; Pred. No. 6.6e-90;
iive 0; Mismatches 0;
                                                                                                                                                                                                                42..45
/note= "Asn is N-glycosylated"
                                                                                                                                                                                 17.19
/note= "Asn is N-glycosylated"
                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 6; Fig 7D; 165pp; English.
                                                                                                                                    Human endogenous retrovirus
                                                                                                                                                                                                                                                                                                                                       98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                      97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                    97EP-0112482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAX07188, AAX07189.
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-143118/12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 153; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                       (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            153 AA;
                                                                                                                                                                                   Modified-site
                                                                                                                                                                                                                    Modified-site
 21-MAY-1999
                                                                                                                                                                                                                                                                   W09905527-A2
                                                                                                                                                                                                                                                                                                                                     22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                      23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                    22-JUL-1997;
                                                                                                                                                                                                                                                                                                      04-FEB-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conrad B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 associated human endogenous retrovirus (IDDMK1.2-22) env protein.

The retrovirus has Superantigen (SAg) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence (without SAg activity but still able to induce an immune response) are useful in vaccines to treat or prevent SAG-related autoimmune disease; nucleic acid sequences encoding (modified) SAg are important in pathogenesis of autoimmune disease, probably by activating autoreactive T cells. The method is very specific (it can differentiate between expressed and non-expressed viral nucleic acids) and can be used corp plasma samples can be tested without extensive preparation and diagnosis can be made before clinical signs are apparent, allowing early intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating
             IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sequence is that of an insulin-dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 840; DB 20;
Pred. No. 6.6e-90;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            85.9%; Scc.
100.0%; Pred
0; F
                                                                                                                                                                                                                                                                                                                                                                                                                                      or preventing autoimmune disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW97745 standard; Protein; 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 31; Fig 7D; 92pp; English.
                                                                                                                                                                                   97EP-0401773
                                                                                                                                                                                                                97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                  WPI; 1999-097928/09.
N-PSDB; AAX07514.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                  Conrad B, Mach B;
                                                                                                                                                                                                                                                (MACH/) MACH B F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153 AA;
                                                                                                                                                                               23-JUL-1997;
```

The

Matches 153;

8 g ò 윤 à

61 121 121

Query Match Best Local

Sequence

ö

Gaps

ö

Length 153;

61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120

ò

AAW97745;

AAW97745 RESULT a

g

1 mvtpvtwmdnpievyvndsvwvpgptddrcpakpeeegmminisigyhyppiclgrapgc 60

```
The sequence is that of an insulin-dependent diabetes mellitus associated human endogenous retrovirus (IDDMK1.2-22) env/fs-sag protein. The retrovirus has Superantigen (SAG) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence (without SAG activity but still able to induce an immune response) are useful in vaccines to treat or prevent SAG-related autoimmne disease; nucleic acid sequences encoding (modified) SAG can be used similarly to treat such diseases. Retroviral-encoded SAG are important in pathogenesis of autoimmune disease, probably by activating autoreactive T cells. The method is very specific (it can differentiate between expressed and non-expressed viral nucleic acids) and can be used even where the pathogen is an ubsquitous endogenous retrovirus. Blood
Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating or preventing autoimmune disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         diagnosis can be made before clinical signs are apparent, allowing early intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             plasma samples can be tested without extensive preparation and
                                                                                                                                                                                                                                                                                                                                IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                             Human endogenous retrovirus IDDMK1.2-22 env/fs-sag protein.
                                                                       121 pkgskntevlvweecvansvvilqnnefgtiid 153
                                                      121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                 AAW95693 standard; Protein; 561 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 31; Fig 7E; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97EP-0401773.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97EP-0401773.
                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-097928/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   561 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MACH/) MACH B F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAX07515.
                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-JUL-1997;
                                                                                                                                                                                                                                                           08-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                EP893691-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27-JAN-1999
                                                                                                                                                                                                                    AAW95693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Seguence
                                                                                      g
                                                                                                                                                                                                     8
                                                      ò
```

```
.nsulin-dependent diabetes mellitus; autoimmune disease; diagnosis;
                                                                                                                                                                                                                                                                                                                                              New isolated human endogenous retrovirus - used to develop product for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Expression of
                                                                                                                            Human endogenous retrovirus IDDKK1.2-22 envelope protein.
                                                                                                                                           HERV; IDDKK1.2-22; superantigen; SAg; antigen; IDDM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              human retroviral SAg can be used as a DNA vaccine.
                                                                                                                                                                                             17.19
/note= "Asn is N-glycosylated"
                                                                                                                                                                                                           42. 45
/note= "Asn is N-glycosylated"
                                                                                                                                                          therapy; vaccine; envelope protein; env gene.
                                    PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                      Location/Qualifiers
                                                                                Ā
                                                                                                                                                                                                                                                                                                                                                                            Claim 6; Fig 7E; 165pp; English.
                                                                               AAW97746 standard; Protein; 561
                                                                                                                                                                         Human endogenous retrovirus.
                                                                                                                                                                                                                                                                              97EP-0401773.
                                                                                                                                                                                                                                                                                     97EP-0112482
                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                 WPI; 1999-143118/12.
                                                                                                                                                                                                                                                                                                    (MEDI-) MEDIGEN SA.
                                                                                                                                                                                             Modified-site
                                                                                                                                                                                                            Modified-site
                                                                                                                                                                                                                                                                                    22-JUL-1997;
                                                                                                             21-MAY-1999
                                                                                                                                                                                                                                  W09905527-A2
                                                                                                                                                                                                                                                               22-JUL-1998;
                                                                                                                                                                                                                                                                             23-JUL-1997;
                                                                                                                                                                                                                                                 04-FEB-1999
                                                                                                                                                                                                                                                                                                                  Conrad B,
                                                                                               AAW97746;
                                                                  φ
              61
61
                            121
                                           121
                                                                                              g
                            ŏ
                                          a
ŏ
```

0

Gaps

0

Indels

85.9%; Score 840; DB 20; 100.0%; Pred. No. 4.2e-89; iive 0; Mismatches 0;

Conservative

Query Match Best Local Similarity Matches 153; Conserv

ò

Length 561;

MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60

SXCCC

ô g ò g ò g

ö

```
The present invention relates to single exon nucleic acid probes (SENP: see AAI31315-AA157546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placents. The probes are useful for antenatal diagnosis of
  specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      by probe for measuring placental gene expression.
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human placenta –
                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                         Probe; microarray; human; placenta; antenatal diagnosis;
                                                                                                                                                              Length
                                                                                                                                                                                                                                                                 47 YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 220; DB 22;
Pred. No. 2.7e-18;
1; Mismatches 7;
                                                                                                                                                           Score 220; DB 22;
Pred. No. 2.7e-18;
                                                                                                                                                                                                              1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 27; SEQ ID No 26727; 654pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Rank
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                             standard; Protein; 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22.5%;
83.3%;
                                                                                                                                                           22.5%;
83:3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0207456.
2000US-0608408.
2000US-0632366.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0236359
2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-JAN-2001; 2001WO-US00663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0234687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                              40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            from human placenta. The
human genetic disorders.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptide #495 encoded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-488897/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Similarity
40; Conserv
                                                                                                                                                           Query Match
Best Local Similarity
                                                                                   AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genetic disorder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200157272-A2.
                                                                                48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-SEP-2000;
27-SEP-2000;
04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JUN-2000;
03-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Matches 40
                                                                                     Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                             AAM26458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAM26458;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Penn SG,
                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                     AAM26458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
     22288
                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME OF COLOR OF STREET OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention relates to human single exon nucleic acid probes (SENP: see AA11068-AA120459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical espithenial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Note: The sequence data for this patent did not form part of the printed
                                                                                                                                                                                                                                                                                                                                                                                                                  LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide #486 encoded by probe for measuring cervical gene expression.
                                                                                                                                                                                                                                                                                       1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        human; microarray; gene expression; cervical epithelial cell;
                                                                                                                                                                                                                                                                                                               T t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells -
                                                                                                                                                                                                                                      ;
0
  the endogenous SAg in IDDM suggests a general model according which self SAg-driven and systemic activation of autoreactive cells leads to organ-specific autoimmune disease.
                                                                                                                                                                                   Length 561;
                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                   85.9%; Score 840; DB 20;
                                                                                                                                                                                                            Pred. No. 4.2e-89;
                                                                                                                                                                                                            Best Local Similarity 100.0%; Pred. No. 4.2
Matches 153; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 27; SEQ ID No 18878; 487pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rank DR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAM14052 standard; Protein; 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0207456.
2000US-0608408.
2000US-0632366.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-JAN-2001; 2001WO-US00670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0236359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0234687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-488901/53.
                                                                                                        561 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cervical cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200157278-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     03-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-SEP-2000;
27-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM14052;
                                                                                                        Sequence
                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Probe;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAM14052
```

ö

ဖ

AAM01792

σ

RESULT

AAM01792

```
The sequences given in AAR29646-51 are encoded by a DNA fragment derived from the Entomopoxvirus Amsacta moore! (AmEPV) around the spheroidin DNA sequence. The open reading frames encode the spheroidin protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                itself and also other structural or regulatory genes associated with spheroidin. Bey Spheroidin has no significant amino acid homology to any previously reported protein. It is a non-essential protein which makes the coding region desirable as a site for the insertion of exogenous DNA. The spheroidin gene is naturally expressed at high levels. Small fragments of the surrounding DNA can be used as regulatory sequences if placed in operative association with foreign
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New viral vectors and chimeric vaccines - comprise entomopoxvirus expression system contg. spheroidin or thymidine kinase sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Entomopoxvirus; non-essential; regulatory sequences; vector.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene expression; vector; insect cell culture; culture; AmEPV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       441 clkpkvpknlrlwgwildcdt----srfikhmadgsddldldvrln----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.2%; Score 80.5; DF 24.3%; Pred. No. 3.8; ative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR55576 standard; Protein; 1003 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Fig 2; 110pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gruidl ME, Hall RL, Moyer RW;
                                                          AmEPV Spheroidin protein G5R.
                                                                                                                                                                                                                                                                                                                                                                                                                                             91US-0657584.
92US-0827685.
                                                                                                                                                                                                                                                                                                                                                                               92WO-US00855.
20-MAY-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (UYFL ) UNIV FLORIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1992-316172/38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AmEPV spheroidin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAQ28301
                                                                                                                                                                                       Amsacta moorei
                                                                                                                                                                                                                                                                                                                                                                               12-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                             19-FEB-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JAN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-JAN-1995
                                                                                                                                                                                                                                               WO9214818-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Spheroidin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mammal cell
                                                                                                                                                                                                                                                                                                                03-SEP-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR55576;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               $\circ{\circ}{\circ}$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\circ$\ci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (see AA10010-AA11006). The present sequence is a peptide encoded by one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probe hybridises at high stringency to a nucleic acid expressed in the human breast. The probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aeticlogy. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative
                                                                                                                                                                                                                                                                                Peptide #474 encoded by probe for measuring human breast gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       breast disease and non-carcinoma tumours. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel single exon nucleic acid probe used to measuring gene expression in a human breast \cdot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present invention relates to novel single exon nucleic acid probes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                       Probe; human; breast disease; breast cancer; development disorder;
inflammatory disease; proliferative breast disease; non-carcinoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      47 YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPR 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 220; DB 22;
Pred. No. 2.7e-18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 27; SEQ ID No 10532; 322pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR29648 standard; Protein; 1003 AA.
                                                                                     AAM01792 standard; Protein; 48 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; 2000US-0180312.
; 2000US-0207456.
; 2000US-0608408.
; 2000US-0632366.
; 2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22.5%;
83.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0236359.
2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29-JAN-2001; 2001WO-US00661
                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Penn SG, Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-476286/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 40; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200157270-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27-SEP-2000;
04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        03-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-AUG-2001
```

ŝ

29;

Indels

39; DB 13;

Seguence

Amsacta moorei entomopoxvirus

WO9413812-A.

AAR29648;

AAR29648

ò g

Length 1003;

```
This sequence is encoded by the Amsacta moorel entomopoxvirus (AmEDV) spheroidin gene, which is an example of the gene of the invention, which encodes a 115 kDa protein. EPV spheroidin and thymidine kinase promoters can be used in DNA constructs and vectors for expression of heterologous genes in insects or mammalian cells, e.g. vectors containing Bacillus thuringiens!s toxin genes for use in insect control, or recombinant vaccinia or swinepox viruses for use as
                                                                                                                                                                                                                                                                                                                                                                       60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Novel expression system for the expression of heterologous sequences in insect and mammalian host cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY30165-70 represent proteins encoded by open reading frames (ORFs)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Spheroidin; Entomopoxvirus; expression system; replication; heterologous gene expression; thymidine kinase; poxvirus; vaccinia; swinepox virus; insect pest control; immunity.
             Entomopoxvirus nucleic acid sequences - used in DNA constructs vectors for expression of heterologous genes in, e.g. insect
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Spheroidin protein encoded by Amsacta moorei entomopoxvirus genome.
                                                                                                                                                                                                                                                                                                        8.2%; Score 80.5; DB 19; Length 1003; 24.3%; Pred. No. 3.8; Live 16; Mismatches 39; Indels 29;
                                                                                                                                                                                                                                                                                                                                                                                             111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Columns 45-50; 55pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Column 61-68; 72pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (UYFL ) UNIV FLORIDA RES FOUND INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Moyer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY30169 standard; Protein; 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Amsacta moorei entomopoxvirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              95US-0544332.
91US-0657584.
92US-0827685.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Υ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95US-0544332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92WO-US00855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92US-0991867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 24.39
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ľį
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hall RL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-457596/38
                                                                                                                                                                                                                                                        1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAZ10081
                                                                                                                                                                                                                          viral vaccines.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-OCT-1995;
19-FEB-1991;
30-JAN-1992;
12-FEB-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-OCT-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US5935777-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gruidl ME,
                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY30169;
                                               cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
               New
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY30169
ò
                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60 CLMPAVQ-----NMLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Entomopoxvirus; spheroidin gene; AmEPV; thymidine kinase; promoter; insect control; viral vaccine.
                                                                                                                                                                                                                       New entomopoxvirus polynucleotide sequences, proteins and vectors - are used for expression of heterologous proteins in both insect and mammalian host cells
                                                                                                                                                                                                                                                                                                                    The sequence of the Amsacta moorel entomopoxvirus spheroidin gene and its flanking regions was determined. The spheroidin gene can be used as the location for the insertion of heterologous DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1003;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                    8.2%; Score 80.5; DB 15; 24.3%; Pred. No. 3.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AmEPV entomopoxvirus spheroidin protein G5R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Pred. No. 3.8; 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                       insect and mammalian expression systems.
                                                                                                                                                                                                                                                                                     Disclosure; Page 68-71; 118pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW41301 standard; Protein; 1003 AA
                                                                                                                                          RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (UYFL ) UNIV FLORIDA RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           91US-0657584.
92US-0827685.
92WO-US00855.
                                                                                                                                            Moyer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             93US-0107755.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Moyer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93US-0107755
                                          93WO-US11907
                                                                        92US-0991867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; AAV14507, AAV14517.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 24.3 tes 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hall RL,
                                                                                                                                          Hall RL,
                                                                                                       (UYFL ) UNIV FLORIDA
                                                                                                                                                                       WPI; 1994-217887/26.
N-PSDB; AAQ66808.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1998-168476/15.
                                                                                                                                                                                                                                                                                                                                                                                                     1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Amsacta moorei
                                        07-DEC-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22-NOV-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-FEB-1991;
30-JAN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22-NOV-1993;
                                                                        07-DEC-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US5721352-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-FEB-1998
          23-JUN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-FEB-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gruidl ME,
                                                                                                                                       Gruidl ME,
                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW41301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
```

AAW41301 RESULT

Best Loc Matches

ò

õ g

ŝ

æ

```
29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Canine von Willebrand factor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99WO-US18153.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (UNMS ) UNIV MICHIGAN STATE.
(VETG-) VETGEN LLC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0132652.
                                                                                                                                                              21.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (UNMI ) UNIV MICHIGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Venta PJ, Brewer GJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-256241/22.
N-PSDB; AAZ51952.
                                                                                                         2813 AA;
                                                                                                                                                               Local Similarity
ses 46; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Canis familiaris.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200009533-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Binding-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Binding-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24 - FEB - 2000
                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY70557;
                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                        75
                                                                                                                                                                                                                                                                                                           131
                                                                                                                                                                             Matches
  a
                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                             à
                                   that is capable of directing the replication and expression of a heterologous gene in a selected host cell. The expression system comparises an EVP promoter sequence operably linked to the selected heterologous gene sequence operably linked to the selected heterologous gene sequence operably linked to the selected heterologous gene sequences and the production of selected capacity in insect and mammalian host cells e.g. human, rodent and proteins in insect and mammalian host cells e.g. human, rodent and capacity insect by thymidine kinase and spheroidin genes can also be used in vertebrate poxyiruses such as vaccinia and swinepox virus. The expression vectors can also be used for the control of insect pests through the insertion of a gene encoding an insect toxin into the quantities of the toxin. Spheroidin and thymidine kinase are nonessential proteins which makes them ideal for the insertion of exogenous DNA and they are capable of operating in a vertebrate poxyirus (e.g. vaccinia) mammalian cell expression vector system. Pox viruses are able to stimulate cell-mediated and humoral immunity.
                                                                                                                                                                                                                                                                                                                                    5;
of the Amsacta moorel entomopoxvirus spheroidin gene and flanking sequences. The DNA is used to make expression systems of the invention. The specification describes an Entomopoxvirus (EPV) expression system
                                                                                                                                                                                                                                                                                                                                                                                       482
                                                                                                                                                                                                                                                                                                                                                           60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This is the canine von Willebrand Factor (vWF) sequence. The encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                von Willebrand Factor; vWF; mutation; von Willebrand disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Canine von Willebrand Factor gene - for developing products to detect carriers of mutation causing canine von Willebrand disease
                                                                                                                                                                                                                                                                                                                                   29;
                                                                                                                                                                                                                                                                                                         Length 1003;
                                                                                                                                                                                                                                                                                                                                                                             111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                               : | | : | : | : | : | | | : | | | | : | : ::| - rndiclkqaikqhytnviileyantypnctlslgnnrfnnvfdmndnkti 532
                                                                                                                                                                                                                                                                                                                                   39; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yuzbasiyan-gurkan V;
                                                                                                                                                                                                                                                                                                          DB 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Canine von Willebrand Factor (vWF) sequence.
                                                                                                                                                                                                                                                                                                         8.2%; Score 80.5; Di
24.3%; Pred. No. 3.8;
                                                                                                                                                                                                                                                                                                                                 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Pages 16-26; 51pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW54347 standard; Protein; 2813 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Venta PJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97WO-US12606.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UNMI ) UNIV MICHIGAN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             960S-0020998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 24.39
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              vWD; detection; breeding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Brewer GJ, Schall WD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1998-120800/11.
                                                                                                                                                                                                                                                                   1003 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAV26525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9803683-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29-JAN-1998
                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Canis sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAW54347;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Canine;
                                                                                                                                                                                                                                                                                                                                                                                                                                         483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 1
AAW54347
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
8
gene is used in a method for detecting a mutation associated with canine von Willebrand disease (vWD) disease, where the mutation is a base deletion at codon 88 of the canine vWF gene. The methods provided in the specification and the products can be used for detecting carriers of the mutation that causes canine vWD. The methods can be used by breeders to reduce the frequency of the disease-causing allele and hence the incidence of the disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVL--- 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---VWEECVANSVVI------LQNNEFGTI------IDLGTSRSILPQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Novel canine von Willebrand factor cDNA sequence and methods for detecting carriers of the defective von Willebrand factor gene which
                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                         26 TDDRCPAKPEEEGMMINIS---IGYHYPPIC-----LGRAPGCLMPAVQNWLVEVPT 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Canine; von Willebrand factor; vWF; von Willebrand's disease; vWD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          detection; diagnostic test; Scottish terrier; Doberman pinscher;
Shetland sheepdog; Poodle; dog.
                                                                                                                                                                                                                                                                                                                                          Length 2813;
                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schall WD;
                                                                                                                                                                                                                                                                                                                                                                                                  73;
                                                                                                                                                                                                                                                                                                                                          DB 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       698.700 7
7label Integrin_binding_site
2507.2509 7label Integrin_binding_site
                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 16; ; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yuzbasiyan-Gurkan V,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | ::: : :||||
| 1031 cadtkkvpldsspavchnnimkqtmvdsscr
                                                                                                                                                                                                                                                                                                                                          8.2%; Score 80.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                164 LLRT------NSVVSKCTSESSCR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY70557 standard; Protein; 2813 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
```

causes von Willebrand's disease in canines -

```
.;
8
                                                                           The present sequence is a canine prepro-von Willebrand factor (VWF) having homology with the human prepro-vWF sequence. Mutation in the vWF gene causes von Willebrand's disease (VWD) in dogs. The present sequence is used in diagnostic tests for detecting carriers of the mutation that causes von Willebrand's disease in canines such as Scottish terriers, Doberman pinschers, Shetland sheepdogs, Manchester terriers and Poodles. Such tests may be used by breeders to reduce the frequency of the disease-causing allele and the incidence of disease. The vWF cDNA sequences may be used to determine the genetic defect that causes vWD in other breeds as well as other species.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           131 --VWEECVANSVVI------LQNNEFGTI------IDLGTSRSILPQ 163
||: :: || : || : ||
971 svvwdhrlsisvtlkrtyqeqvcglcgnfdgiqnndftssslqieedpvdfgnswkvnpg 1030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   75 VSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVL---- 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 63; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26 TDDRCPAKPEEEGMMINIS---IGYHYPPIC-----LGRAPGCLMPAVQNWLVEVPT 74
                                                                                                                                                                                                                                                                                                                                                                                    DB 21; Length 2813;
                                                                                                                                                                                                                                                                                                                                                                                                                            73;
                                                                                                                                                                                                                                                                                                                                                                              Query Match 8.2%; Score 80.5; DB Best Local Similarity 21.8%; Pred. No. 16; Matches 46; Conservative 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1031 cadtkvpldsspavchnnimkqtmvdsscr 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               164 LLRT-----NSVVSKCTSESSCR 181
                                          Example 1; Page 46-55; 62pp; English.
                                                                                                                                                                                                                                                                                                                      2813 AA;
                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

Search completed: April 9, 2002, 16:56:45 Job time: 479 sec

	-	

	A STATE OF THE STA	
	en e	
	9	· c
		e - 4),
	*	
43		
		•
	*	
	· ·	
		*
		. ×.
		•
		F.
		Tr

```
09HDC0
071037
071072
09GMP3
09B047
086521
09W524
09W54
09W54
09W54
09W54
            7
11
11
10
10
10
10
10
10
                                                                                                                       10
 152
124
206
306
340
347
226
633
633
973
295
220
221
222
223
224
226
228
332
332
                                                                          (without alignments)
373.734 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                    US-09-490-700-40
978
1 MVTPVTWMDNPIEVYVNDSV......PQLLRTNSVVSKCTSESSCR 181
                                                               April 9, 2002, 17:06:33; Search time 70.84 Seconds
                                                                                                                                                                                        473505
         GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                    473505 seqs, 146272329 residues
                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                              OM protein - protein search, using sw model
                                                                                                                                                   Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
5: sp_human:*
5: sp_human:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                          sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                       sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                              sp_rodent:*
                                                                                                                                                                                                                                                                         SPTREMBL_17:*
                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                     10:
11:
12:
14:
                                                                                                     Title:
Perfect score:
                                                                                                                                          Scoring table:
                                                                                                                        Sequence:
                                                                                                                                                                    Searched:
                                                                                                                                                                                                                                                                         Database
                                                                Run on:
```

	Description	· 095280 homo sapien	O42043 unclassifie	Q9qc06 human endog	Q9unw3 homo sapien		homod	рошо	уошо	homo	095281 homo sapien	Q9yna9 human endog	Q9ubu4 homo sapien	Q69384 human endog	Q69386 human endog	Q9yna7 human endog	095282 homo sapien	Q9yna5 human endog	Q9ukh7 homo sapien	Q9ukh3 homo sapien
SUMMARIES	ID	095280	042043	900060	Q9UNW3	Q9UNW2	Q9HDB8	014273	ОЭПКНЭ	095284	095281	Q9YNA9	Q9UBU4	Q69384	069386	Q9YNA7	095282	Q9YNA5	д9икн7	09икнз
	DB	4	12	12	4	4	4	7	4	4	4	12	4	12	12	12	4	12	4	4
	Query Match Length DB	153	153	260	153	153	245	1361	2294	153	153	694	669	669	514	694	153	694	277	698
950	Ouery Match	85.9	85.9	85.0	82.9	82.3	82.3	81.0	80.7	80.3	80.0	79.7	79.7	79.7	79.6	9.6	79.4	79.4	79.1	76.9
	Score	840	840	831	811	802	802	792	789	784	782	779	779	779	778	778	777	777	774	752.5
	Result No.	1	71	m	4	'n	Q	7	60	6	10	11	12	13	14	15	16	17	18	19

9	095283 homo sapien Q9hdco homo sapien 071037 human endog	Q99mp3 macaca fasc Q9du47 chiba virus O80770 arabidopsis	008521 cricetulus Q9tuq3 sus scrofa Q30489 equus cabal	Q9vt54 drosophila Q9r194 staphylococ Q48348 enterococcu	O22158 arabidopsis O9xxi5 caenorhabdi	Q9xx16 caenorhabdi Q9f960 arabidopsis	Q9spml lycopersico Q9rla8 mus musculu	Oyp/y8 schizosacch Q55633 synechocyst	Q564U/ 11STERIA MO Q9C9b4 arabidopsis P893107 vam mosaic	Q9n533 caenorhabdi Q9xa09 streptomyce	ENTS		153 AA.	ed) sequence update) annotation update)		1	a; Vertebrata; Euteleostoml; ini; Hominidae; Homo.			O.Y., Vargas A., Rao J.,	diab			endogenous retrovirus gene, a , and identification of a CA repeat		ank∕DDBJ databases.		2B36E1656DF0C01A CRC64;	ore 840; DB 4; Length 153; ed. No. 5.6e-79; Mismatches 0; Indels 0; Gaps 0;	KPEEEGMMINISIGYHYPPICLGRAPGC 60		LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
005283	095283 Q9HDC0 071037	Q9GMP3 Q9DU47 Q80770	008521 Q9TUQ3 Q30489	Q9VT54 Q9RL94 Q48348	022158 09XXI5	Q9XXI6 Q9FG60	09SPM1 09R1A8	Q55633	Q5640/ Q9C9B4 P89307	533 A09	ALIGNMENT		PRT;	Created) Last sequ			; Craniata; V ; Catarrhini;		8243;	Chen				of an end abetes, a		EMBL/GenBank/DDBJ 	.·	1; 2B36E1	SCC P1	GPTDDRCPA		HMVSGMSLR
-	4 4 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	12	111 6 7	5 7 7	10	10	11	n (1)	10	502			ARY;	10,	Ì	1	Chordata; Primates;		3-977	tant rrv R	ediat			tion 1 di	1	the 3.1;	. T :	05 MW	85.9%; 100.0%; ive (SVWVF	SVWVF	SRFTY
15.2	154 124 206 206	300 1785 990	347 843 226	633 972 973	492	514	711	318	971 623 459	3885			PRELIMINARY	(TrEMBLrel. (TrEMBLrel.					PubMed	Course	mune-me	. (0	. X . ;	caliza r type	us.";	99) to AC6889:	AD3305	; 173	ch 1 Similarity 100 153; Conservative	EVYVND	EVYVND	PTVSPN
76.8	63.5 50.2	12.0 8.9 8.5	8 8 8 7 9 4		88.1	8.7 9.9	, , , 8. %.	7.8	7.6	7.6					· [s (Human)	Metazo utheri 9606:		OM N.A 49692;	lason A	and im	661)01	OM N.A Jinno	and lo ene fo	ts loc	MAR-19 864; A	984; A otein.	153 AA	ilarit Conse	WMDNPI	 MMDNPI	NWLVEV
_	0 /51.5 1 621 2 491 3 472	4100	78 6	210	м 4 1	വരാ	~ & c	n 0 ,	+ C1 F	4. N		1 1	095280	01-MAY-1999 01-MAY-1999 01-MAY-1999	TVELOPE	Homo sapiens	Eukaryota; Metazoa; Mammalia; Eutheria; NCBI TaxID=9606;	[1]	SEQUENCE FROM N.A. MEDLINE-98449692; PubMed-9778243	Lan M.S., Mason A., Coutant R.	"HERV-K10s	[2]	SECUENCE FR Hasuike S.,	"Isolation candidate q	marker at i	Submitted (EMBL; AF084 EMBL; AF124	EMBL; AF134 Envelope pr	SEQUENCE 153 AA; 17305 MW; 2B36E10	, Mat Loca les	1 MVTPVT	 1 MVTPVT	61 LMPAVQ
C	7 00 00	000	0 0 0	ммм	mmi	יחרחנ	ייטריי	. . .	444	44		RESUL	OJC OJI	70 10 10 10 10 10	D E	SOS	388	R.	R R X	RA	RT	N K	R R	RT	R.	DR G	K K	SOS	Query Best Match	Qy	QQ	δλ

~

ö

Gaps

ö

Length 560;

9 9

042043; 042043

121

g δ g

```
61 LMPAVQNWLVEVPTVSPNSRFTYHWVSGMSLRPRVNCLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                         LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                              1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                              1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
 Score 831; DB 12;
Pred. No. 2.2e-77;
                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                     PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 PKGSKNTEVLVWEECVANSAVILQNSEFGTIID 153
                                                                                                                                                                                                     121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                PRT;
 85.0%;
99.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13,
                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ENVELOPE PROTEIN RIC-7
                                                                                                                                                                                                                                                                                                                                                                                                      sapiens (Human)
               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 148; Conserv
                 Best Local Sim
Matches 152;
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9UNW2
Q9UNW2;
                                                                                                                                                                                                                                                                                                             09UNW3;
                                                                                                                                                                                                                                                                                             Q9UNW3
                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                        HOMOH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENV.
                                                                                                                                                                                                                                                              RESULT
Q9UNW3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09UNW2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                             ŏ
                                                                                         a
                                                                                                                         ŏ
                                                                                                                                                                                     δ
                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=99445825; PubMed=10516026;
Toenjes R.R. Czauderna F., Kurth R.;
"Genome wide screening, clonomy chromosomal assignment and expression
of full-length human endogenous retrovirus type K (HERV-K).";
J. Virol. 73:9187-9195(1999).
EMBL; Y18890; CAB56604.1; ". 8A4565663901BC3A CRC64;
61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                     MEDIINE-97386455; PubMed-9244304;
Conrad B., Weissmahr R.N., Boni J., Arcari R., Schupbach J., Mach B.
"A human endogenous retroviral superantigen as candidate autoimmune
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                               2B36E1656DF0C01A CRC64;
                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          85.9%; Score 840; DB 12;
100.0%; Pred. No. 5.6e-79;
ive 0; M1smatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      560 AA.
                                                              PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             153
                                            121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID
                                                                                                                                                                                                                                                                              Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBL_TaxID=45617;
                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AF012337; AAC58457.1; -. EMBL; AF012336; AAC58456.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            153 AA; 17305 MW;
                                                                                                                                                                                    05,
05,
08,
                                                                                                                                                                                                                                                                                                                                                                                    gene in type I diabetes.";
Cell 90:303-313(1997).
                                                                                                                                                                                                                                                              unclassified Retroviridae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 85.9
Best Local Similarity 100.
Matches 153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                01-JAN-1998 (TrEMBLrel.
01-JAN-1998 (TrEMBLrel.
01-NOV-1998 (TrEMBLrel.
ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-2000 (TrEMBLrel.
01-MAY-2000 (TrEMBLrel.
01-MAR-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                             NCBI_TaxID=35276;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Envelope protein.
SEQUENCE 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENV PROTEIN.
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 LMPAVQNWLVEVPTVSPISRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDHCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                           Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                  .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 153;
                                                                                                                                                                                                                                       Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Butele
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606,
                                                                                                                                                                                                                                                                                                                           [1]
SEQUENCE FROM N.A.
MEDLINE-98449692; PubMed-9778243;
Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A., I Gomez R., Chalew S., Garry R., Maclaren N.K.;
"HERV-KIOS and immune-mediated (type 1) diabetes.";
cell 95:14-16(1998).
EMBL; AF084655; AAC68894.1; -.
EMBL; AF084655; AAC68894.1; -.
ENELOPE PLOTEIN.
SEQUENCE 153 AA; 17226 MW; 4632F483BFD9517A CRC64;
                                                01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
11-MAY-2000 (TEMBLrel. 13, Last annotation update)
ENVELODE PROTEIN RIC-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82.9%; Score 811; DB 4;
96.7%; Pred. No. 5.5e-76;
Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A
```

900060

121

g ò g ò

ò

ç

ò g ð a ð g

```
TRANSCRIPTASE).
SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
SIMILARITY: THE PROTEASE BELONGS TO THE PEPTIDASE FAMILY ALSO
KNOWN AS THE EUKARYOTIC ASPARTYL PROTEASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                         retrovirus genome related
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Many human endogenous retrovirus K (HERV-K) proviruses are unique
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606;
                                                                                        Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-99400989; Pubmed-10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1361;
                                                                                     Craniata; Vertebrata; Eutele
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 792; DB 4; Length 13
Pred. No. 6.5e-73;
1; Mismatches 7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CEB91B3F407B9498 CRC64;
                                                                                                                                                                                                                                                      MEDLINE=87036922; PubMed=3021993;

MOON M., Yasunaga T., Miyata T., Ushikubo H.;

Nucleotide sequence of human endogenous retrovirus genc

the mouse mammary tumor virus genome.";

J. Virol. 60:589-598(1986).

-! SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE

TRANSCRITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Curr. Biol. 9:861-868(1999).
-!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-UNN-2001 (TrEMBLrel. 17, Last annotation update)
GAG-PRO-POL-ENV PROTEIN.
   Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2294 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL, M14123, AAA88033.1; -...
InterPro, IPR001037; Integrase_C.
InterPro, IPR0021308; Integrase_an.
InterPro, IPR002156; RNaseH.
InterPro, IPR00477; RVF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA-directed DNA polymerase.
SEQUENCE 1361 AA; 153797 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PP00552; integrase; 1. Pfam; PF02022; Integrase_Zn; 1. Pfam; PF00075; rnaseH; 1.
                                                                            Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.0%;
94.8%;
01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 145; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF00078; rvt; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00665; rve; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                         Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                                    TISSUE-LIVER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lenz J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        humans."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09UKH9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           о9икн9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             о9икн9
DDE DREAD DR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 LMPAVQNWLVEVPTVSPNSRFTYHWVSGMSLRPRVNYLQDFSYQRSLKFRPKGKDCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIVYRYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jinno Y., Sugimoto J.;
"Human endogenous retrovirus HERV-K(II) and flanking sequences.";
Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AB047240; BAB11760.1; -.
SEQUENCE 245 AA; 27904 MW; 198F26D65ED56DDB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
0
                                                                                                                                                                                                                                                                                                                                                                             ;
                                                      J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82.3%; Score 805; DB 4; Length 245; 96.7%; Pred. No. 4e-75; Live 0; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                   Length 153;
                                                         Rao
                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                         Vargas A.,
                                                                                                                                                                                                                              153 AA; 17406 MW; 90B91BCE0040C5A3 CRC64;
                  MEDLINE-98449692; PubMed-9778243;
Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A. Gomez R., Chalew S., Garry R., Maclaren N.K.;
HERV-K10s and immune-mediated (type 1) diabetes.";
Cell 95:14-16(1998).
EMBL, AF084870; AAC68899.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
                                                                                                                                                                                                                                                                                                                Query Match 82.3%; Score 805; DB 4; L
Best Local Similarity 96.7%; Pred. No. 2.3e-75;
Matches 148; Conservative 0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  245 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 PKESKNTEVLVWEECVANSAVILONNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 PKESKNTEVLVWEECVANSAVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01, Created)
01, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 82.3
Best Local Similarity 96.7
Matches 148; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel.
01-NOV-1996 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                   Envelope protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                              SEQUENCE
```

Q9HDB8;

09HDB8

о9нрв

ö

ö

S

014273;

014273 ID 01 AC 01 DT 01

014273

RESULT

õ g ò g ð 8

```
Envelope protein.
SEQUENCE 153 AA; 17391 MW; 4088880E06FB7677 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                694 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 PKESKNTEVLVWEECVDNSAVILQNSEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; . 153 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last seq
01-MAY-1999 (TrEMBLrel. 10, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Human endogenous retrovirus K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 80.0%;
Best Local Similarity 94.1%;
Matches 144; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10,
                                                                                                               Query Match
Best Local Similarity 94.1;
Matches 144; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-1999 (TrEMBLrel.
ENVELOPE PROTEIN RIC-3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Envelope protein.
SEQUENCE 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENV PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  095281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09YNA9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9YNA9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9YNA9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ID DDT TO DD TO DD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S S
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1795 MVTPVTWMDNPIEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGYRYPPICLGRAPGC 1854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human).
Sukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rao J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80.7%; Score 789; DB 4; Length 2294; 94.8%; Pred. No. 2.5e-72; ive 1; Mismatches 7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART; SM00343; ZnF_C2HC; 2.
SMOSITE; PS00141; ASP_RROTEASE; 1.
PROSITE; PS00143, ASP_ROTETROY; 1.
ASPARTYI PROLEASE; Hydrolase; RNA-directed DNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2294 AA; 257099 MW; 425C81FDDF8A80D0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A.,
Gomez R., Chalew S., Garry R., Maclaren N.K.;
"HERV-K10s and immune-mediated (type 1) diabetes.";
Cell 95:14-16(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1915 PKESKNTEVLVWEECVDNSAVILQNNEFGTIID 1947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   153 AA
                                                                                    Asp_prot_retrov duTPase.
                                                                                                                                                              InterPro; IPR000721; Gag_p24.
InterPro; IPR000467; G_patch.
InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_zn.
                                                 InterPro; IPR001969; Asp_protease
                                                                                                                                                                                                                                                                                                                                                                                                                                              PF005037; Gag_pl0; 1.
PF00607; gag_p24; 1.
PF00552; integrase; 1.
PF02002; Integrase_zn; 1.
PF00075; rnaseH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
EMBL; AF164611; AAD51793.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF084869; AAC68898.1; -.
                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000477; RVTse.
InterPro; IPR001878; Znf_CCHC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00078; rvt; 1.
Pfam; PF00098; zf-CCHC; 2.
ProDom; PD000946; dUTPase; 1.
                                                                                                                                            Gag_p10.
                                                                                                                                                                                                                                                                          InterPro; IPR002156; RNaseH.
InterPro; IPR001584; Rve.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, ENVELOPE PROTEIN RIC-6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SM00443; G-patch; 1
SM00343; ZnF C2HC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 94.8
Matches 145; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                            PF00692; dUTPase;
PF01585; G-patch;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF00665; rve; 1
PF00077; rvp; 1
                                                                                                               InterPro; IPR001428;
                                                                                                                                      InterPro; IPR003322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Aspartyl pro
Zinc-finger
                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             095284
```

121

RESULT
095284
1D 009
AC 099
AC 099
DT 011
DT 011
DT 011
DT 01
DT 0

61

ò 셤 ò g

a

SMART;

Pfam;

Pfam; Pfam;

ON WEEK WAY OF A PROPERTY OF A

Pfam; Pfam; Pfam; Pfam; Pfam; Pfam;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ÷
                                                                                                                                                                                                            LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                       61 LMPAVQNWLVEVPTVSPTSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKKI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 LMPAVQNWLVEVPIVSPISRFTYHWVSGMSLRPRVNYLQDFPYQRSLKFRPKGKPCPKEI 120
                                                         Gaps
                                                                                                       1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Euteleostomi;
                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-98449692; PubMed-9778243;
Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A., Rao J., Gomez R., Chalew S., Garry R., Maclaren N.K.;
"HERV-Kilos and immune-mediated (type 1) diabetes.";
Cell 95:14-16(1998).
EMBL; AF084866; AAC68895.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 153;
     Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 153 AA; 17360 MW; 28C01AD348C6D806 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E
Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
80.2%; Score 784; DB 4; I 94.1%; Pred. No. 3.4e-73; Live 2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80.0%; Score 782; DB 4; L
94.1%; Pred. No. 5.4e-73;
tive 0; Mismatches 9;
```

ö

OC NAR RAP OC SOR RET RAP DR RTT RAP DR RTT

õ

ò g ð g

```
61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                           1 MVIPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-95074858; PubMed-7983704;
MEDLINE-95074858; PubMed-7983704;
Loewer R., Toenjes R.R., Korbmacher C., Kurth R., Loewer J.;
Loewer R., Toenjes R.R., Korbmacher C., Kurth R., Loewer J.;
"Identification of a Rev-related protein by analysis of spliced
transcripts of the human endogenous retroviruses HTDV/HERV-K.";
J. VIROL. 69:141-149(1995).
EMBL: X82272; CAA5732.1;
EMBL: X82272; CAA5732.1;
SEQUENCE 699 AA; 79217 MW; 3844DA6DB0D76DFC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 699;
   Length 699;
                                 9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 779; DB 12; I
Pred. No. 6.6e-72;
2; Mismatches 9;
 Score 779; DB 4;
Pred. No. 6.6e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
MEDLINE-97213939; Pubmed-9060628;
Tonjes R.R., Limbach C., Lower R., Kurth R.;
                                 2; Mismatches
                                                                                                                                                                                                         121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    79.7%; Score 779; 92.8%; Pred. No. 6.
                                                                                                                                                                                                                           13 PKESKNTEVLVWEECVANSAVILQNNEFGTIID
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human endogenous retrovirus.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=11827;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 79.7%;
Best Local Similarity 92.8%;
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01,
01,
08,
                                                                                                                                                                                                                                                                                                                                                              01,
01,
08,
                 Best Local Similarity 92.8
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                               (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POL/ENV GENE (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                            01-NOV-1996 (
01-NOV-1996 (
01-NOV-1998 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1996
01-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                              ENV MRNA.
 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POL/ENV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        069386
                                                                                                                                                                                                                                                                                                                              069384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14
                                                                                                                                                                                                                                                                                                            069384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ID
DT
DT
DT
DT
DT
DT
DT
RX
RX
                                                                                                     g
                                                                                                                                                                       q
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                       ò
                                                                                                                                       ò
                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  t
                                                                                                       and
                                                                                                                                                                                                                                                                                                             61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'An almost-intact human endogenous retrovirus K on human chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lenz J.;
"Many human endogenous retrovirus K (HERV-K) proviruses are unique
                                                                Toenjes R.R., Czauderna F., Kurth R.;
"Full-length human endogenous retrovirus type K (HERV-K) elements
encoding Gag, Pol and Env proteins are localised on chromosomes 7
                                                                                                                                                                                                                                                                                                                                                                                 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        desse E.; Further characterization of the almost intact human endogenous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-99178257; Pubmed-10080172;
Mayer J., Sauter M., Racz A., Scherer D., Mueller-Lantzsch N.
                                                                                                                                                                                                                         Length 694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Reus K., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalla; Eutheria; Frimates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases EMBL; AF164614; AAD51798.1; -. EMBL; AF074086; AAF81168.1; -. EMBL; AF074086; AAD51098.1; -. EMBL; AF074086; AAD51098.1; -.
                                                                                                                                       to the EMBL/GenBank/DDBJ databases
                                                                                                                                                         380.1; -.
78690 MW; 82F91825669CF25B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              699 AA; 79217 MW; 3844DA6D80D76DFC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                      79.7%; Score 779; DB 12;
92.8%; Pred. No. 6.5e-72;
Live 2; Mismatches 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  227 PKESKNTEVLVWEECVANSAVILQNNEFGTIID 259
                                                                                                                                                                                                                                                                                                                                                                                                                              121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        669
Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE=99400989; PubMed=10469592;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              retrovirus K on human chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genet. 21:257-258(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Curr. Biol. 9:861-868(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY 2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAR-2001 (TrEMBLrel. 16, ENVELOPE PROTEIN.
                                                                                                                                                                                                                      Query Match
Best Local Similarity 92.88
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                      EMBL; Y17832; CAA76880.
SEQUENCE 694 AA; 786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human).
                                                                                                                                       Submitted (JUL-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Envelope protein.
SEQUENCE 699 AA
              NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9UBU4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Meese E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                humans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      090BU4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12
                                                                                                                                                                                                                                                                                                                                                                                              167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
Q9UBU4
```

ö

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.

Toenjes R.R., Czauderna F., Kurth R.;
"Full-length human endogenous retrovirus type K (HERV-K) elements
encoding Gag, Pol and Env proteins are localised on chromosomes 7 and
                                                                                                                                                                                                                                                                                       61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                               61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGXHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                     .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                     79.6%; Score 778; DB 12; Length 514; 92.8%; Pred. No. 5.8e-72; tive 2; Mismatches 9; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79.6%; Score 778; DB 12; Length 694; 92.8%; Pred. No. 8.3e-72; ive 2; Mismatches 9; Indels (
"Expression of human endogenous retrovirus type K envelope glycoprotein in insect and mammalian cells."; J. Virol. 71:2747-2756(1997).
                                                                                 Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases. EMBL; x92887; CAA63481.1; -NOV TER 545 SEQUENCE 514 AA; 58200 WW; 86410F35B82DF231 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases EMBL; Y17833; CAA76883.1; -. SEQUENCE 694 AA; 78692 MW; 498A1D745366D0EE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                 694 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                              121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                               Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
VGEL_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-2000 (TrEMBLrel. 13, ENV PROTEIN.
                                                                                                                                                                     Query Match 79.6'
Best Local Similarity 92.8
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 79.6
Best Local Similarity 92.8
Matches 142; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                       SEQUENCE FROM N.A.
                                                                     Toenjes R.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9YNA7;
                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9YNA7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    227
                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
Q9YNA7
 qq
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
```

Search completed: April 9, 2002, 17:06:34 Job time: 583 sec

```
GenCore version 4.5
```

Copyright (c) 1993 - 2000 Compugen Ltd.	OM protein - protein search, using sw model	April 9, 2002, 17:07:04 ; Search time 23.68 Seconds (without alignments) 280.251 Million cell updates/sec	US-09-490-700-40 e: 978 l MytpytmmDnpievYvnDsVPQLLRTNSVVSKCTSESSCR 181	e: BLOSUM62
J	Ж protein - protei	Run on: Ap	Title: US Perfect score: 97 Sequence: 1	Scoring table: BI

100059 Total number of hits satisfying chosen parameters:

100059 seqs, 36664827 residues

Searched:

Gapop 10.0 , Gapext 0.5

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_39:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		Description	P10267 homo sapien		094616 schizosacch	homo		5 amsac	Q28295 canis famil		P11828 glycine max	P23512 rattus norv	O9qz67 mus musculu	P38931 saccharomyc	Q63921 rattus norv	P16950 t amylopull	P10259 mouse mamma				018388 drosophila		Q00400 dugesia tig	-		_	_	Q53515 rhizobium t	P03374 mouse mamma	002061 schizosacch	P56546 mus musculu	P55317 homo sapien	m	P25244 bacteriopha	P49748 homo sapien
SUMMARIES		ID	ENV1_HUMAN	RAI2_MOUSE	RPC5_SCHPO	RAI2_HUMAN	S24A_HUMAN	SPHR_AMEPV	VWF_CANFA	CYRB_MOUSE	GLC3_SOYBN	HN3A_RAT	P2CD_MOUSE	SRB9_YEAST	PGH1_RAT	APU_THETY	ENV_MMTVB	PM5P_HUMAN	APU_THEET	SCN1_SCHPO	IMB_DROME	VL2_HPV54	HMH1_DUGTI	TRPE_ACICA	MASY_PICAN	TREF_ECOLI	COMA_BACSU	NODU_RHITR	ENV_MMTVG	RPB2_SCHPO	CTB2_MOUSE	HN3A_HUMAN	VGA_BPAL3	VGA_BPPHK	ACDV_HUMAN
		8 :	П	-		7	-	-	-		-	-	-	Н	-	-	-	П	-	Н	-	7	-		-	-	-	٦	-		-	-	Н		-
	:	Match Length	584	529	348	530	1078	1002	2813	968	481	466	298	1420	602	1475	688	1222	1481	335	884	470	533	497	555	549	214	575	688	1210	445	473	494	464	655
đ	Query	Match	78.6	4.6	8.5	8.4			7.8		٠	7.6	7.5	7.5	7.4	٠	7.4	7.3	7.3	٠	7.3	•	7.2	7.1		7.0			٠				6.9		6.9
	,	Score	768.5	91.5		82.5	82.5	80.5	76.5	97	75		73.5	73.5	72.5	72.5	72	71.5	71.5	71	71	70.5	70.5	69.2	69.5	68.5	68	99	68	68	67.5	67.5	67.5	67.5	67.5
	Result	. ; 0 ;		7	e	4	S	9	7	8	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	58	30	31	32	33

529 AA.

PRT;

STANDARD;

RESULT 2 RAI2_MOUSE : ID RAI2_MOUSE AC Q9QVY8;

q οy qq

c 3 5 0 0 5 7 5 C C 8							Ä		
P26675 drosophila O9nqx4 homo sapien P22437 mus musculu D89029 mus musculu P40522 saccharomyc P53242 saccharomyc P35446 rattus norv Q09457 caenorhabdi P13705 mus musculu P13611 homo sapien O43010 schizosacch				Euteleostomi; Homo.	genome related to	POTENTIAL). (POTENTIAL).	h 584; s 1; Gaps	LGRAPGCLMP 63 LGRAPGCLMP 60	KTCPKEIPKG 123 KPCPKEIPKE 120
				Eut.		9999999999	Length Indels	PPIC	RPKG RPKG
	rs	4 AA.	odate) update)	Craniata; Vertebrata; Eutele Catarrhini; Hominidae; Homo.	rovirus	YPTOTEIN. LINKED (GLCNAC) (PLINKED (GLCNAC) (OB 1; -68; 6;	PVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMP 	AVQNWLVEVPTVSPNSRETYHMVSGMSLRPRVNYLQDESYQRSLKERPKGKTCPKEIPKG
DROME HUMAN MOUSE MOUSE TYEAST TYEAST CAEEL CAEEL MOUSE HUMAN SCHPO	ALIGNMENT	584	pdn no	ata; rhin	ikub enou	tein ED (CED (CED (CED (CED (CED (CED (CED (C	ore 768.5; ed. No. 1.4 Mismatches	PEEB 	RVNY
	ALIG	PRT;	0, Created) 0, Last sequence update) 0, Last annotation updat	Crani	93; , Ushikubo endogenous genome.";	Polyprotein N-LINKED	8 8	RCPAK 	MSLRP
SOS. MY56 MTN. MTN. Y1FC YG20 YG20 YO3PC MSH.			Created) Last seq Last ann POLYPRO		=30219 ata T. human virus 6).	3		PTDC PIDC	MVSG MVSG
		;	Creat Last Last V POLN). Chordata; Primates;	ed=3 Lyat Lyat hu c vi 986)	otein 4 4 9 9 0 7 7 6 6 5 9 9 9	78.68 94.78 ive	VPG 	TYH
1595 1742 602 624 624 640 292 807 963 1091 3396 466		STANDARD;		n). ; Chc ; Pri	PubMe	716040406070	vat	NDSVA - - - NDSIV	PNSRE PISRE
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		STA	(Rel. (Rel. (Rel. ELATE	s (Huma Metazoa Utheria 9606;	36922; Pu unaga T., sequence ammary tu 0:589-598		78 ilarity 94 Conservativ	NPIEVYV 	VEVPTVS
67.5 67.5 67.5 66.5 66.5 66.5 66.5 66.5		T 1 HUMAN ENVI_HUMAN	01-MAR-1989 01-MAR-1989 01-MAR-1989	ENV. Homo sapiens (Human) Fukaryota; Metazoa; Mammalia; Eutheria; NCBI_TaxID=9606;	SEQUENCE FROM N.A. MEDLINE-87036922; PubMed=3021993; Ono M., Yasunaga T., Miyata T., Ushikubo H.; Nuoleotide sequence of human endogenous ret the mouse mammary tumor virus genome."; J. Virol. 60:589-598(1986).	GARBOHYD CARBOHYD CARBOHYD CARBOHYD CARBOHYD CARBOHYD CARBOHYD 3 CARBOHYD 3 CARBOHYD 4 CARBOHYD 4 CARBOHYD 4 CARBOHYD 4 CARBOHYD 5 SEQUENCE 58	Query Match Best Local Similarity Matches 142; Conser	4 PVTWMD	64 AVQNWL 61 AVQNWL
. 444444444444444444444444444444444444		RESULT 1 ENV1_HUMAN ID ENV1_	70008	ENV HOMC EUK Mamr	ME SE	866666666666666666666666666666666666666	Query M Best Lo Matches		
		RES ENV ID	4 F F F F F F	8 0 0 0 0 S	RP RA RT RL DR	KW FT FT FT FT SO	Ome	Oy Dp	Qy Dp

~

```
13
                                                                                                                                             Nogi
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                           ۷;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Shpakovski G.V., Shematorova E.K.; "Molecular cloning and characteristics of rpc19+ and rpc40+ Schizosaccharomyces pombe genes, coding for common subunits of nuclear RNA polymerase I and III."; Bloorg. Khim. 24:933-937(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
DNA-DIRECTED RNA POLYMERASES I AND III 40 KDA POLYPEPTIDE (EC 2.7.7.6)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LMPAVQNWLVEVPTVSPNSRFTYHWVS-GMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90 MPIHMQVEGSSAPELNPNGNATYVMTTQGPVQLPVVLEQHVFQHLNS-----PLVLPQE 143
                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32 LITTEAWNINSTDLVKKALVTVPGPSILNPPAE-SQSGMALKVAATV-LQPLCLGESPVV 89
                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                     Jonk L.J.C., de Jonge M.E., Vervaart J.M., Wissink S., Kruijer W. "Isolation and developmental expression of retinoic-acid-induced
                                                                                                                                                                                                                                                                                                                                                                                                                                         29;
                                                                                                                                                                                                                                                                                                                                                                                                                Length 529;
                                                                                                                                                                                                                                                                                                                                                                                                                                        73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120 IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLL 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----SEAQPQLL 169
                                                                                                                                                                                                                                                                                                                                                                          529 AA; 57178 MW; 10AA48B170FCDBD0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                              9.4%; Score 91.5; DB 1;
25.3%; Pred. No. 0.17;
tive 22; Mismatches 73;
20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2011 (Rel. 40, Last annotation update)
RETINOIC ACID-INDUCED PROTEIN 2 (3F8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             348
                                                                                                                                                                                                                                                                                                                                                              PRO-RICH
                                                                                                                                                                                                                                                                                                                                     EMBL; X76652; -; NOT_ANNOTATED_CDS. MGD; MGI:1344378; Rai2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-972;
MEDLINE-99179688; Pubmed-10079952;
                                                                                                                            SEQUENCE FROM N.A.
MEDLINE-94148137; PubMed-8314004;
                                                                                                                                                                                            Dev. Biol. 161:604-614(1994).
-!- INDUCTION: BY RETINOIC ACID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RPC40 OR RPA42 OR SPBC1289.07C.
                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match 9.4's
Best Local Similarity 25.3's
Matches 42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Schizosaccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                   NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RPC5_SCHPO
                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              144 AP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         094616;
                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                genes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
RPC5_SCHPO
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. The European Bioinformatics Institute. The European Bioinformatics Institution as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 55 AND TRNA GENES.
SIMILARITY: BELONGS TO THE RPB3/RPC5 RNA POLYMERASE SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wood V., Rajandream M.A., Barrell B.G., Oliver K., Harris D.;
Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     124 SKNTEV-----LVWEECVANSVVILQNNEFGTIIDLGTSRS------ILPQLLR 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71 EVPTV----SPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFR---PKGKTCPKEIPKG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               272 KKQARVADVRKDTVSREC-----LRHPEFADKVQLGRVRDHYLFSVESTGIMKPDVLF 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    52; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 15 DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS SHARED BETWEEN YEAST RNA POLYMERASES I AND III. THE AC19 AND AC40 SUBUNITS INTERACT WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EVYVNDSVWVP-GPTDDRCPAKPEEEGMMINISIGYHYPPICLGR-APGCLMPAVQNWLV 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE
                                                                                                                                                 Shpakovski G.V., Shematorova E.K.; "Rpc19 and Rpc40, two alpha-like subunits shared by nuclear RNA polymerases I and III, are interchangeable between the fission and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.5%; Score 83; DB 1; Length 348; 23.5%; Pred. No. 0.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    70; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001514; RNA_polD_30KD.
ProDom; PD002883; RNA_polD_30KD; 1.
PROSITE; PS00446; RNA_POL_D_30KD; 1.
Transferase; DNA-directed RNA polymerase; Transcription;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           348 AA; 39154 MW; A9C209362521E4AE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 23.5%; Pred. No. 0.7;
Matches 46; Conservative 28; Mismatches
                                                                                                          MEDLINE=20009333; PubMed=10541858;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBCELLULAR LOCATION: NUCLEAR.
                                                                                                                                                                                                                                                                                                      Genet. 36:208-214(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF082512; AAD44503.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AB021859; BAA77385.1; -
EMBL; AL035675; CAB38687.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       167 TNSVV---SKCTSESS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              325 IKSIAVLKSKCLAVKS 340
[2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                        Imazawa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nuclear protein.
SEQUENCE 348 A
                                                                                                                                                                                                                                                                   budding yeasts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBSTRATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RNA(N)
```

11;

m

```
118791 MW;
                                                                                                                                                                                                                                                                                                     EMBL; AJ131244; CAA10334.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  23.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 23.89
Matches 39; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                        440
                                                                                                                                                                                                                                                                                                                                                                                  1078 AA;
                                                                                                                                                                                                                                                                                                                                     Multigene family.
                                                                                                                            LYMPHOCYTES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SPHR_AMEPV
                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P29815;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
SPHR_AMEPV
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROTEIN TRANSPORT PROTEIN SEC24A (SEC24-RELATED PROTEIN A) (FRAGMENT).
                                                                                                                                                                                                                                                                    MEDLINE-99168896; PubMed-10049581; Walpole S.M., Hiriyana K.T., Nicolaou A., Bingham E.L., Durham J., Vaudin M., Ross M.T., Yates J.R., Sieving P.A., Trump D.; Identification and characterization of the human homologue (RAI2) of a mouse retinoic acid-induced gene in Xp22."; Genomics 55:275-283(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 LMPAVQNWLVEVPTVSPNSRFTYHMVS-GMSLRPRVNYLQDFSYQRSLKFRPKGKTC 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pagano A., Letourneur F., Garcia-Estefania D., Carpentier J.-L., Orci L., Paccaud J.-P.; Sec24 proteins and sorting at the endoplasmic reticulum."; J. Biol. Chem. 274:7833-7840(1999).
                                                                                                                                                                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .,
m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 530;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9879EE869DC6188F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8.4%; Score 82.5; DB 1; 25.6%; Pred. No. 1.3;
                                                                                        20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
RETINOIC ACID-INDUCED PROTEIN 2.
                                                            530 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1078 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRO-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-99175155; PubMed-10075675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE 530 AA; 57148 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20-AUG-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; AF136587; AAD33688.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 Similarity 25.6% 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                            STANDARD;
                                                                                                                                                                         Homo sapiens (Human)
                                                                                                                                                                                                                                         [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE-B-cell;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MIM; 300217;
                                                         RAI2_HUMAN
Q9Y5P3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S24A_HUMAN
095486;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                           RAI2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S24A_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90
```

g

ò

g

å

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modifited and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7;
VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO THE GOLGI APPARATUS. COPII ACTS IN THE CYTOPLASM TO PROMOTE THE TRANSPORT OF SECRETORY, PLASMA MEMBRANE, AND VACUOLAR PROTEINS SUBUNT: COPII IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC13/31 COMPLEX, THE SEC13/31 COMPLEX AND SARI.

SUBGELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR (BY SIMILARITY).

TISSUE SPECIFICITY: EXPRESSED IN FIBROBLASTS, HEPATOCYTES, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86 VSGMSLRPR----VNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSV- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  323 MSGLSLQPEGLRVVNLLQERNMLPSTPLKPPVPNLHEDIQKLNCNPELF---RCTLTSIP 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      272 SVGYSYP-----SLPPGYQNTTPPGATGVPPSSLNYPSGPQAFTQTPLGANHLTTS 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Amsacta moorei entomopoxvirus (AmEPV).
Viruses; dSDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
Entomopoxvirus B.
NCBI_TaxID=28321;
                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hall R.L., Moyer R.W.;
"Identification, cloning, and sequencing of a fragment of Amsacta moorei entromopovarius DNA containing the spheroidin gene and three vaccinia virus-related open reading frames.";
J. Virol. 65:6516-6527(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=92185464; PubMed=1545219;
Banville M., Dumas F., Trifiro S., Arif B., Richardson C.;
"The predicted amino acid sequence of the spheroidin protein from Amsacta moorei entomopoxvirus: lack of homology between major occlusion body proteins of different poxviruses.";
J. Gen. Virol. 73:559-566(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AJ131244; CAA10334.1; -.
Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.4%; Score 82.5; DB 1; Length 1078; 23.8%; Pred. No. 3.1; Ive 23; Mismatches 61; Indels 41.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZINC FINGER-LIKE.
W; 9A3214F52D28FD3F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             141 ---VILONNEFGTIIDLGTSRSILPQLLRTNSVVSKCTSESSCR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         380 QTQALLNKAKLPLGLLLHPFKDLVQLPVVTSSTIVRC---RSCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01.FEB-1996 (Rel. 33, Last annotation update)
SPHEROIDIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1002 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. MEDLINE=92046310; Pubmed=1942245;
```

us-09-490-700-40.rsp

```
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                               HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                     EMBL;
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                     CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                               -- 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Venta P.J., Li J., Yuzbasiyan-Gurkan V., Brewer G.J., Schall W.D.; "Complete sequence of the structural gene for canine von Willebrand factor and identification of a mutation causing Scottish terrier von Willebrand's disease.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
"The canine von Willebrand factor gene: sequence and expression of
a region encoding the glycoprotein ID/IX binding domain.";
Submitted (JAN-1994) to the EMEL/GenBank/DDBJ databases.
-i- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMBOSTAGIS, IT
PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Stoy S.J., Shibuya H., Nonneman D.J., Holzhauer J., Mohammed I.H., Johnson G.S.;
         PROFECTIVE COMPLEXES IN WHICH THE VIRIONS ARE EMBEDDED IN THE CYTOPLASM OF THEIR INSECT HOSTS.
SUBUNIT: MAY FORM DISULFIDE BONDS LINKED AGGREGATES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
                                                                                                                                                                                                                                                                                                             29;
                                                                                                                                                                                                                                                                                    DB 1; Length 1002;
                                                                                                                                                                                                                                                                                                                                                              440 CLKPKVPKNLRLWGWILDCDT----SRFIKHMADGSDDLDLDVRLN------
 FUNCTION: MAJOR COMPONENT OF VIRAL OCCLUSION BODIES, THE
                                                                                                                                                                                                                                                                                                                                                                                       111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                       : | |: | : |: :: | 482 -RNDICLKQAIKQHYTNVIILEYANTYPNCTLSLGNNRFNNVFDMNDNKTI 531
                                                                                                                                                                                                                                                                                                             39; Indels
                                                                                                                                                                                                                                 1 ACETYLATION.
114738 MW; DC0D61B157645F85 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Montgomery R.R., Fahs S., Montgomery M.W.;
Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                         PIR; JQ1436; PYVZAM.
Viral occlusion body; Acetylation; Late protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VWF_CANFA STANDARD; PRT; 2813 AA. 028295; Q28311; Q9TS14; 01-NOV-1997 (Rel. 35, Created) 20-NOC-2001 (Rel. 40, Last sequence update) 20-NOC-2001 (Rel. 40, Last annotation update) VON WILLEBRAND FACTOR PRECURSOR (VWF).
                                                                                                                                                                                                                                                                                  8.2%; Score 80.5; DE
ilarity 24.3%; Pred. No. 4.5;
Conservative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1234-1669 FROM N.A.
                                                                                                                                                                EMBL; M75889; AAA42378.1; -.
                                                                                                                                                                                 EMBL; M77182; AAA42383.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    familiaris (Dog)
                                                                                                                                                                                                                                             1002 AA;
                                                                                                                                                                                                                                                                                                Local Similarity
les 27; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IISSUE-Blood;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F8VWF OR VWF.
                                                                                                                                                                                                                  INIT_MET
MOD_RES
SEQUENCE
                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Canis
                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                     9
Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                   INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
--- SIMILARITY: CONTAINS 3 VWFA DOMAINS.
--- SIMILARITY: CONTAINS 3 VWFC DOMAINS.
--- SIMILARITY: CONTAINS 4 VWFD DOMAINS.
--- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
--- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                        -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Endothelial cell; Repeat; Cell adhesion; Signal.
1 22 BY SIMILARITY.
23 763 BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VON WILLEBRAND FACTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SITE
SITE
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELL ATTACHMENT
CELL ATTACHMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELL ATTACHMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AMINO-TERMINAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
               VASCULAR INJURY (BY SIMILARITY).
SUBUNIT: MULTIMERIC (BY SIMILARITY).
TISSUE SPECIFICITY: BLOOD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VWFD 3.
VWFA 1.
VWFA 2.
VWFA 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               E2.
VWFC 1.
VWFC 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VWFD 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VWFC
                                                                                                                                                                                                                                                                                                                                                                                          L16903; AAA30903.1; -. AF099154; AAD04919.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000359; Cys_knot.
                                                                                                                                                                                                                                                                                                                                                                                                                               U66246; AAB93766.1; -. P04275; 1ATZ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE; PS01185; CTCK_1; 1. PROSITE; PS01225; CTCK_2; 1. PROSITE; PS50234; VWFA; 3. PROSITE; PS01208; VWFC; 3.
                                                                                                                                                                                                                                                                                                                                                                          EMBL; L76227; AAB05549.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VWFADOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001846; vwd.
InterPro; IPR002035; vwFA.
Pfam; PF00007; Cys_Knct; 1.
Pfam; PF01826; TIL; 4.
Pfam; PF00092; vwa; 3.
Pfam; PF00093; vwa; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SMART; SM00041; CT; 1.
SMART; SM0327; VWA; 3.
SMART; SM00214; VWC; 5.
SMART; SM00216; VWD; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  853
1013
1453
1453
1665
1871
2102
2261
2326
2326
2495
2495
2650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR000004;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR001007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PR00453;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2216
2255
2429
2580
2724
531
698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             764
788
826
867
867
1277
1498
1691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Plasma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
```

```
RA GORMOLD D.M., ILON N., KITAMULAT T., Schreurs J., Yonehara S.,

RA GORMOLD D.M., ILON N., KITAMULAT T., Schreurs J., Yonehara S.,

RA GORMOLD M., ILON N., KITAMULAT T., Schreurs J., Yonehara S.,

RA Yahara I., Arai K., Miyajima A.;

RA Yahara I., Arai K., Miyajima A.;

RI ike protein: identification of another member of the cytokine

RT ike protein: identification of another member of the cytokine

RT ike protein: identification of another member of the cytokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: identification of another member of the stokine

RT ike protein: Asia identification of another member of the stokine

RT ike protein: Asia identification of another member of the stokine

RT ike protein: Asia identification of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of the stokine

RT ill identification of another member of the stokine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 VPG-PTDDRCPAKPEEEGMMINISIG-YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNS 79
                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
CYTOKINE RECEPTOR COMMON BETA CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY.

BY SIMILARITY.

N-LINKED (GLCNAC. . .) (POTENTIAL)

N-LINKED (GLCNAC. . .) (POTENTIAL)

N-LINKED (GLCNAC. . .) (POTENTIAL)

8CE16EDFDC07a999 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80 RFTYHMVSGMSLR------PRVNYLQDFSYQRSLKFRP----KGKTC 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.8%; Score 76; DB 1; Length 896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FIBRONECTIN TYPE-III 1. FIBRONECTIN TYPE-III 2.
                                                 01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
10-AXY-2000 (Rel. 39, Last annotation update)
CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR
CSF2RB OR CSFR2B1 OR AIC2B OR IL3RB1.
                   896 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MGD; MGI:1339759; Csf2rbl.
InterPro; IPR002996; CRACk_rcptor_2.
InterPro; IPR002985; Cytok_rcptor_2.
InterPro; IPR001777; FN_III.
InterPro; IPR003531; Hematopo_rcptor_S_F1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00041; fn3; 2.
PROSITE; PS01355; HEMATOPO_REC_S_F1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99111 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M34397; AAA37204.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                     Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             141
350
896 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 29; Conserv
                                                                                                                                                                                                           NCBI_TaxID=10090;
                   CYRB_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                     P26955
                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      131 --VWEECVANSVVI------LDLGTSRSILPQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TDHVCDATCSAIGMAHYLTFDGLKYLFPGECQYVLVQDYCGSNPGTFRILVGNEGCSYPS 918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      75 VSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVL---- 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches 74; Indels 63; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26 TDDRCPAKPEEEGMMINIS---IGYHYPPIC-----LGRAPGCLMPAVQNWLVEVPT 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7.8%; Score 76.5; DB 1; Length 2813; 21.3%; Pred. No. 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5DF93E1E5E72F80C CRC64;
                                                                                                                                                                                                                                                         1942 (BY SIMILARITY). SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          I -> V (IN REF. 2).

WCH -> ICQ (IN REF. 2).

3 -> A (IN REF. 1).

F -> L (IN REF. 2).

L -> P (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
               SIMILARITY.
                                                                                                                                                                                                      SIMILARITY.
SIMILARITY.
SIMILARITY.
1942 (BY SI
                                                                                                                                                                                                                                                                                                  SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1031 CADTKKVPLDSSPAVCHNNIMKQTMVDSSCR 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       164 LLRT------NSVVSKCTSESSCR 181
BY SIMILAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match /.us;
Best Local Similarity 21.3%;
Matches 45; Conservative ?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA; 309716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1231
1515
1574
2223
2290
2357
2400
2546
2585
2790
                                                                                                                                                                                                                                                              1940
2123
2085
2088
2001
22774
2774
2778
2804
2806
2806
2801
290
156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2406
                                                                                                                                                                                                                                                              1899
1972
1950
1927
1993
2724
2739
2750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1231
1515
1515
1523
2223
2357
2357
2357
2357
2357
2585
2790
2585
               DISULFID
                                                    DISULFID
                                                                                          DISULFID
                                                                                                                                               DISULFID
                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                            DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                   DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 859
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

7;

Gaps

26;

846 E-TEDLVQDLSVKKFPYQPMPQAPAIQFFKSLKHQDYLSLPPWDNSQSGKVC 896

g

RESULT

us-09-490-700-40.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5,
                                                                                                                                             Glycine max (Soybean).
Bukaryota, Vilidiplantae, Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudlcotyledons; core eudlcots; Rosidae;
eurosids I; Rabales; Fabaceae, Papliionoideae; Phaseoleae; Glycine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STINAPQEIYIQQGSGIFGMIFPGCPSTFEEPQQKGQSSRPQDRHQKIYHFREGDLIAVP 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45 IGYHYPPICLGRAPGCLMPAVQNWLV---EVPTVSPNSRFTYHMVSGMSLRPRVNYLQDF 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------WMYNNEDTPVVAVSLIDTNSFQNQLDQMPRRFYLAGN 184
                                                                                                                                                                                                                                                                                                                                           62; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TWMDNPIEVYVNDSVWVPGPTDDRCPAKPEE------EGMMINIS 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: BELONGS TO THE 11S SEED STORAGE PROTEIN (GLOBULINS)
                                                       01-OCT-1989 (Rel. 12, Created)
01-OCT-1989 (Rel. 12, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
GLYCININ G3 PRECURSOR (CONTAINS: GLYCININ A SUBUNIT; GLYCININ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.7%; Score 75; DB 1; Length 481; 18.2%; Pred. No. 6.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INTERCHAIN (BY SIMILARITY).
5F3C3148DF6241A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Multigene family; Signal.
                              481 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A SUBUNIT.
B SUBUNIT.
                                                                                                                                                                                                                                                                           Cho T.-J., Nielsen N.C.;
"The glycinin Gy3 gene from soybean.";
Nucleic Acids Res. 17:4388-4388(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000459; Seedstore_11s.
Pfan; PF00190; Seedstore_11s; 1.
PRNTYS; PR00439; 11SGLOBULIN.
PROSITE; PS00305; 11S_SEED_STORAGE; 1.
                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
STRAIN-CV. DARE; TISSUE-Leaf;
MEDLINE-89296500; PubMed-2740231;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54241 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X15123; CAA33217.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 18.28
Matches 35; Conservative
                              STANDARD;
                                                                                                                                                                                                                                                                                                                                    DISCUSSION OF SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Seed storage protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22
297
477
107
481 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       143 TGFAY-----
                                                                                                                                                                                                      NCBI_TaxID=3847;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FAMILY.
                            GLC3_SOYBN
P11828;
                                                                                                                SUBUNIT].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
RESULT 9
GLC3_SOYBN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83
                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ď
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
ع
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -i- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
102 SYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSIL 161
                                                                                                                                                                                                                                                                                                                                                        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                   -----GFA 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           322 PGPAASPQTLDHSGATATGGGSELKSPASSSAPPISSGPGGWICTPLSPTWLAPHESQLH 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23 PGP-----TDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNWLVE----- 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUCH AS AFP, ALBUMIN, TYROSINE AMINOTRANSFERASE, PEPCK, ETC.

"INTERACTS WITH THE CIS-ACTING REGULATORY REGIONS OF THESE GENES.

-!- SUBCELLULAR LOCATION: NUCLEAR.

-!- TISSUE SPECIFICITY: LIVER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nuclear protein; Transcription regulation; Activator.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "HNF-3A, a hepatocyte-enriched transcription factor of novel structure is requiated transcriptionally."; Genes Dev. 4:1427-1436(1990).
-I- FUNCTION: TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LIVER GENES
                                                                                                                                                                                                                                                                                                                            Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---VPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   382 LKGAPHYSFNHPFSINNL--MSSSEQQHKLDFKAYEQALQYSPYGATLPASLPIG 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.6%; Score 74; DB 1; Length 466; 25.2%; Pred. No. 7.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                    [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 313-337 AND 413-434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Lai E., Prezioso V.R., Smith E., Litvin O., Costa R.H., Darnell J.E. Jr.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06555BA74BD9B7DC CRC64;
                                                                                                                                                                                                                                    01-NOV-1991 (Rel. 20, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
HUF3A OR TCF3A OR TCF-3A.
                                                                                                                                                                                                    466 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FORK-HEAD
                                 185 QEQEFLQYQPQKQQGGTQSQKGKRQQE---
                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=91032994; PubMed=2227418;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00657; FORK_HEAD_1; 1. PROSITE; PS00658; FORK_HEAD_2; 1. PROSITE; PS50039; FORK_HEAD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; A36674; A36674.
TRANSFAC: T00331; -
InterPro; IPR001766; Fork_head.
Pfam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            48774 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X55955; CAA39418.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM00339; FH; 1
                                                                       162 POLLRTNSVVSK 173
                                                                                                         225 PEFLEHAFVVDR 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            466 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 29; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA_BIND
SEQUENCE
                                                                                                                                                                                                    HN3A_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              72
                                                                                                                                                                                                    g
                                Db
                                                                       ŏ
                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
466 TLE-ENCPKALTLRIHDSLNNTLSVGLIPTNSTNTIMDQKNLKMSTPGQMKAQEV 519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                         RESULT 12
SRB9_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
g
                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10;
                                                                                                                                                                                 20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROTEIN PHOSPHATASE 2C DELTA ISOSORM (EC 3.1.3.16) (PP2C-DELTA) (P53-INDUCED PROTEIN PHOSPHATASE 1) (PROTEIN PHOSPHATASE MAGNESIUM-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89 MSLRPRVNYLQDFSYQ-----SKLKFRPKGKTCPKEIPKG-----SKNTE 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PG-----LVEVPTVSPNSRFTYHMVSG 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 TPVTWMDNPIEVYV---NDSVW--VPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRA 57
                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----ANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSV 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 598;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4DB70B5D48539435 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase; Magnesium; Manganese; Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22.1%; Pred. w.
tive 26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.5%; Score 73.5; F
                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR003589; PP2C_catalytic.
InterPro; IPR001932; PP2C_domain.
InterPro; IPR001932; PP2C_sig.
Pfam: PF00481; PP2C; 1.
                                                                                                                                                                                                                                                                                                                                                                          EMBL; AF200464; AAF09251.1; -.
                                                    (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65722 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00332; PP2Cc; 1.
SMART; SM00331; PP2C_SIG; 1.
PROSITE; PS01032; PP2C; 1.
                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000222; PP2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                  MGD; MGI:1858214; Ppmld
                                                                                                                   Mus musculus (Mouse)
                                                                                                DEPENDENT 1 DELTA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      598 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 52; Conserv
                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     129 VLVWEECV----
                                                                                                                                              NCBI_TaxID-10090;
                                                                                                                                                                           STRAIN-129/SVJ;
                                                                                                          PPM1D OR WIP1
                                                   20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                  P2CD_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                         P2CD_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   901
                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

<u>::</u>

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A., Dietrich F.S., Mulligan J., Chenry J.M., Chung E., Duncan M., Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Law H., Lin D., Mosedale D., Nakahara K., Komp C., Lashkari D., Petel F.X., Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N., Winant A., Yelton M., Botstein D., Davis R.W.;
Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.

FUNCTION: COMPONENT OF THE RNA POLYMERASE II HOLOENZYME AND THE MIDDIATON OF ACTIVATION SECONDED.
                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-95293223; PubMed-7774808; Hengartner C.J., Thompson C.M., Zhang J., Chao D.M., Liao S.-M., Koleske A.J., Okamura S., Young R.A.; Association of an activator with an RNA polymerase II holoenzyme."; Genes Dev. 9:897-910(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                       Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 1420;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7F6CF4BBE0FAC918 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yuryev A., Corden J.L.;
Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                   Saccharomyces cerevisiae (Baker's yeast).
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Sacch
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID-4932;
                                                      01-FEB-1995 (Rel. 31, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
01-CT-1996 (Rel. 34, Last annotation update)
SUPPRESSOR OF RNA POLYMERASE B SRB9 (SCA1 PROTEIN).
SRB9 OR SCA1 OR SSN2 OR YDR443C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-GLU.
POLY-SER.
POLY-LEU.
POLY-LEU.
D -> E (IN REF. 2).
E -> V (IN REF. 2).
T -> S (IN REF. 2).
VK -> GE (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          T -> P (IN REF. 2).
Y -> S (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
         1420 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7.5%; Score 73.5; 23.0%; Pred. No. 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; U23812; AAA91316.1; -.
EMBL; U09176; AAA18614.1; -.
EMBL; U33007; AAB64875.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      160000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47; Conservative
         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        664
816
1008
1136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; S48538; S48538.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSFAC; T02153; -. SGD; S0002851; SSN2. Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEOUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-S288C
SRB9_YEAST
P38931;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
```

```
Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGH1_RAT STANDARD; PRT; 602 AA.

Q63921; Q62731; Q63684;

15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROSPAGIANDIN G/H SYNTHASE 1 PRECURSOR (EC 1.14.99.1) (CYCLOOXYGENASE -1) (COX-1) (PROSTAGLANDIN-ENDOPEROXIDE SYNTHASE 1) (PGH SYNTHASE 1) (PGH SYNTHASE 1) (PGH SYNTHASE 1) (PGH SCATAGLANDIN PTGSI OR COX1 OR COX-1.
                                                                                                                                                                                                                                                   CATALYTIC ACTIVITY: ARACHIDONATE + AH(2) + 2 O(2) - PROSTAGLANDIN
                                                                                                                                                     69 INDDPVPTPPAIEHKPVSPDKIGTFTADYSKPNLPPHYALFLKALRRKIYINLALGSHNK 128
                                                                                                                        PICLGRAPGCL-MPAVQNWLVEV-PTVSPNSRFTYHM-VSGMSLRP--RVNYLQDFSYQR 105
                                                                                                                                                                                                                           SLKFRPKG-----KTCPKE--IPKGSKNTEVLVWEECVANSVVILQNNE----FGTIIDL 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Métazoa; Chórdata; Cranlata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-SPRAGUE-DAWLEY;
MEDLINE-94095619; PubMed-8274023;
MEDLINE-94095619; PubMed-8274023;
Wilson C.B., Hwang D.; Tang W.W., Chanmugam P., Soyoola E.,
Wilson C.B., Hwang D.;
"Cloning two isoforms of rat cyclooxygenase: differential regulation of their expression."
Arch. Biochem. Biophys. 307:361-368(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MICROSOMAL MEMBRANE.
-1- MISCELLANEOUS: THIS ENZYME ACTS BOTH AS A DIOXYGENASE AND AS A
                       ---MMINISIGYHYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -i- MISCELLANEOUS: THIS ENZYME IS THE TARGET OF NONSTEROIDAL ANTI-INFLAMMATORY DRUGS SUCH AS ASPIRIN.
-i- SIMILARITY: CONYAINS I EGE-LIKE DOMAIN.
-i- SIMILARITY: BELONGS TO THE PROSTAGLANDIN G/H SYNTHASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  \mathrm{H2} + A + \mathrm{H(2)O}. PATHWAY: FIRST STEP IN THE FORMATION OF PROSTAGLANDINS AND
                                                                                                                                                                                                                                                                                                                          155 G-----TSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                         246 GHLNGHIPTIASYLIPLLEAKKLV 269
                    16 VNDSVWVPGPTDDRCPAKPEEEG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   THROMBOXANES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PEROXIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLS.
                                                                                                                                                                                                                         106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PGH1_RAT |
110 PGH1_RAT |
110 PGH1_RAT |
110 PGH1_PGH |
111 PGH1_PGH |
111 PGH1_PGH |
112 PROS |
113 PROS |
114 PGH |
115 PGH |
115 PGH |
116 PGH |
117 PGH |
117 PGH |
118 PGH |
118 PGH |
118 PGH |
119 PGH |
119 PGH |
111 PGH 
                                                                    8
                                                                                                                                                             qq
                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                     ò
```

```
WLRS--SLRPSPSFTHFLLTHGYWIWEFVNATFIREVLMRLVITVRSNLIPSPPTYNTAH 136
                                                                                                                                                                                                                   CYCLOOXYGENASE (BY SIMILARITY).
PROXIMAL HEME LIGAND (BY SIMILARITY).
ASPIRIN-ACETYLATED SERINE.
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).
N-SI (IN REF. 1).
RL -> GW (IN REF. 1).
C -> A (IN REF. 1).
C -> L (IN REF. 1).
V -> L (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLG---RAPGCLMPAVQN 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 PPPVLLTDA-GVPSPVNPCCYYPCQNQGVCVRFGLD-HYQCDCTRTGYSGPNCTIPEIWT 78
                                                                                                 PROSTIE; PS00022; EGF_1; FALSE_NEG.
PROSTIE; PS01086; EGF_2; FALSE_NEG.
Oxidoreductase; Dioxygenase; Peroxidase; Glycoprotéin; Acetylation;
Prostaglandin biosynthesis; Heme; Iron; Signal; Membrane;
EGF-like domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-AUG-1990 (Rel. 15, Created)
01-AUG-1990 (Rel. 15, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
AMYLOPULLULANASE PRECURSOR (ALPHA-AMYLASE/PULLULANASE) [INCLUDES:
ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE);
PULLULANASE (EC 3.2.1.41) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE);
ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               49;
                                                                                                                                                                                             EGF-LIKE.
DISTAL HISTIDINE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 602;
                                                                                                                                                                   BY SIMILARITY.
PROSTAGLANDIN G/H SYNTHASE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB29400).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100 DF-----SYQRSLKFRPKGKTCPKEIPKGSKNTEVL 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                    N KEF. 1)

REF. 1)

REF. 1)

N REF. 1

N REF. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WLVEVPTVSPNSRFTYHM------VSGMSLR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 72.5; DB
Pred. No. 15;
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 1475 AA
                                  InterPro; IPR002007; Anim_peroxdse.
InterPro; IPR001561; EGF-like.
InterPro; IPR001536; Peroxdse_3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7.4%; Sco
21.7%; Pre
vative 20;
                                                                          PRINTS; PR00457; ANPEROXIDASE.
SMART; SM00181; EGF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              69032
         S67721; AAB29400.2;
U18060; AAA85823.1;
EMBL; U03388; AAA03465.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ¥.
                                                                                                                                                                                                                                                                                       Best Local Sim
Matches 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APU_THETY
P16950;
                                                                                                                                                                                 CHAIN
DOMAIN
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                  MOD_RES
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                     BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
APU_THETY
ID AC P169F
DT 01-7
DT 20-7
DE AV
DE AV
DE AV
GN
                                                                                                                                                                     SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         137
```

```
1080 NAVYNKVEGVEISKDWTWVDAQYVGDSGNNDKYMAKFVPDMVGTWEYIMRFSSNOGHDWT 1139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77176 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M15122; AAA46544.1; -.
PIR; D26795; VCMVMM.
Coat protein; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                               STANDARD;
                                                 1140 YTKG----PDGKT 1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98
474
688
127
143
297
498
                          YORSLKFRPKGKT 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99
475
127
143
297
498
557
688 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Sim
Matches 25;
                                                                                                               ENV_MMTVB
P10259;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROPEP
                                                                                     RESULT 15
ENV_MMTVB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                          103
                                                 a
                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
 q
                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                               KNOWN AS THE ALPHA-AMYLASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1031 THVIGVNNPVEVYA--EVWAQGLTD-----KPGQGENMI-AQLGYRYIGDTVGDA---YY 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---WL-VEVPTVSP---NSRFTYHMVSGM-----SLRPRVNYLQDFS 102
                                                                                                                                                                                                                          "Purification and some properties of the extracellular alpha-amylase-pullulanase produced by Clostridium thermohydrosulfuricum."; Biochem. J. 250:813-818(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 TPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLM 62
                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
                                                                                                                                                                                                                                                                                                              "Proposed acquisition of an animal protein domain by bacteria.";
Proc. Natl. Acad. Sci. U.S.A. 89:8990-8994(1992).
-!- CATALTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN OLIGOSACCHARLDED POLYSACCHARIDES.
-!- CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYZES (1-6).ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00060; FN3; 2.
Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;
                                                                                                            Melasniemi H., Paloheimo M., Hemioe L.;
"Nucleotide sequence of the alpha-amylase-pullulanase gene from
Clostridium thermohydrosulfuricum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7.4%; Score 72.5; DB 1; Length 1475; 28.6%; Pred. No. 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY.
3476C414110BE376 CRC64;
Thermoanaerobacter thermohydrosulfuricus (Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AMYLOPULLULANASE.
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
BY SIMILARITY.
BY SIMILARITY.
                        Bacteria; Firmicutes; Bacillus/Clostridium group; Thermoanaerobacter group; Thermoanaerobacter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14; Mismatches
                                                                                                                                                  Gen. Microbiol. 136:447-454(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A44765, A44765.
InterPro: IPR000461; Alpha_amylase.
InterPro: IPR001077; FN_III.
Pfam; PF00128; alpha-amylase; 1.
Pfam; PF00041; fn3; 2.
                                                                                                                                                                                     STRAIN=E101-69;
MEDLINE-88268757; PubMed-3260488;
                                                                                                                                                                                                                                                                                        PubMed=1409594;
                                                                                     STRAIN=E101-69;
MEDLINE=90362027; PubMed=2391488;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BY
MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M28471; AAA23205.1; -.
                                                                                                                                                                                                                                                                          FIBRONECTIN TYPE III DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           165631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                        MEDLINE=93028390; Pubmed
Bork P., Doolittle R.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1475
1016
1255
629
              :hermohydrosulfuricum).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            enzyme
                                                                                                                                                                                                                                                                                                                                                                                            FORM MALTOTRIOSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 38; Conserv
                                                                          FROM N.A.
                                                                                                                                                                           SEQUENCE OF 32-39.
                                                 NCBI_TaxID=1516;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Multifunctional
                                                                                                                                                                                                                Melasniemi H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63 PAVQN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
SEQUENCE
                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NI WWO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          113 PVGWGSTDPIRVLTNQTMYLGG-----SPDFHGFR-NMSGNVHFEGKSDTLPICLSF 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PVTW-MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYP-----PICL-- 54
                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-STIL2944; PubMed=3027377;
Moore R., Dixon M., Smith R., Peters G., Dickson C.;
"Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of gag and pol.";
J. Virol. 61:480-490(1987).
                                                                   01-MAR-1989 (Rel. 10, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN (CONTAINS: COAT PROTEIN GP52; COAT PROTEIN GP36).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC. . .) (POTENTIAL)
N-LINKED (GLCNAC. . . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.4%; Score 72; DB 1; Length 688; 29.1%; Pred. No. 20; ive 10; Mismatches 31; Indels
                                                                                                                                                                                                                                              Mouse mammary tumor virus (strain BR6).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus.
NCBL_TaxID=11758;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  46CAAC8D61FFADC2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COAT PROTEIN GP52.
688 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LEADER PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search completed: April 9, 2002, 17:07:05 Job time: 569 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SESTPTGCEQVDKQVFLSDTPTVDNN 189
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --GRAPGCLMPAVQNWLVEVPTVSPN 78
```

ŝ

		,			
					2
					4.
	30 ×	* ·			
	72×1	- "	A		
	a				
•					1.0
		-			
	The second of th			r,	
	(s., ,			9 (E)	- 1
				£	(t) **
	* ************************************				*
			ج ن ي. در د		, ai Lie.
		r t		¥ 1 3 4 4	
				-	
	W	y a	200		
	•		*		
	*				
÷v o				and the state of t	
	e e e e e e e e e e e e e e e e e e e				£ -
					•
·					
		· * · · · · · · · · · · · · · · · · · ·		tar J	
		0.			
			, v*		
*.					
		*	E	*.	
					*
				* g	
				ā.	
		· · · · · · · · · · · · · · · · · · ·			
*					
• /•					
	a				· · · · · · · · · · · · · · · · · · ·
		24.	*		

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

_	
mode]	
×	
using sw model	
search,	
protein	
•	
OM protein	
ŏ	

9, 2002, 16:58:14 ; Search time 39.41 Seconds (without alignments) 349.850 Million cell updates/sec April Run on:

US-09-490-700-40 978 1 MVFPVTWMDNPIEVYVNDSV......PQLLRTNSVVSKCTSESSCR 181 Title: Perfect score: Sequence:

219241 segs, 76174552 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

Scoring table:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

1: pirl:* 2: pir2:* 3: pir3:* 4: pir4:* PIR_68:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	 elated	nscrip	RNA p	histo	tnpA -	recurs	domai	prote	prote	prote	eptor	prote	- soyb	orecur	se 1 -	ein F2	n fact	orane	prote	- yea	ü	prote	prote	prote	orane	H/S u	e (EC	proto
Description	retrovirus-related	probable transcrip	DNA-directed	MHC class II	transposase	spheroidin precurs	probable AP2	hypothetical	hypothetical	hypothetical	cytokine receptor	hypothetical	qlycinin G3	glycinin G3 precur	cyclooxygenase 1	unknown protein F2	transcription fact	probable membranė	hypothetical	SRB9 protein	vacuolar sort	hypothetical	hypothetical	hypothetical	probable membrane	prostaglandin G/H	alpha-amylase	bring that i an 1
	! ! ! ! ! !																											
ΩI	VCHUER	T02309	T39358	A53273	JQ0971	PYVZAM	T00399	T26502	T26501	T50303	A35782	S76295	S04605	S11003	S39782	н96766	A36674	T36724	T21236	B57062	T04895	S76888	C86647	T00396	S64909	S69198	A44765	00000
DB	4	~	~	7	~	-	~	7	7	~	ч	~	~	~	~	~	~	~	7	7	~	7	7	7	7	7	~	ď
Length	584	066	348	226	973	1003	295	492	514	704	968	318	481	484	602	623	466	297	467	1420	626	1072	365	461	583	602	1475	400
% Query Match	78.6	8.5	8.5	8.4	8.3	8.2	8.1	8.1	8.1	7.8	7.8	7.7	7.7	7.7	7.6	7.6	7.6	7.5	7.5	7.5	7.5	7.5	7.4	7.4	7.4	7.4	7.4	
Score	768.5	83.5	83	82	81	80.5	79	79	79	97	76	75.5	75	75	74.5	74.5	74	73.5	ص	73.5	73	73	72.5	72.5	72.5	72.5	72.5	33
Result No.	-	7	m	4	Ŋ	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	٥

env polyprotein -	hypothetical prote	Pm5 protein - huma	hypothetical prote	amylopullulanase p	molybdate-binding	scnl protein - fis	conserved hypothet	homeotic protein D	probable transmemb	probable polygalac	hypothetical prote	env polyprotein -	hypothetical prote	hypothetical prote	hypothetical prote
VCMVMM	ce2083	S21977	T25168	828669	B83413	B55164	F83201	S33701	T51150	B86155	S51404	S26388	T33067	T39745	T51502
ч	7	7	~	7	7	7	~	~	~	7	7	7	7	7	7
688	848	1190	1262	1481	251	387	295	533	256	491	570	688	414	433	488
7.4	7.4	7.3	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2	7.1	7.1	7.1
	72 7.4													69.5 7.1	

ALIGNMENTS

Н	
RESULT	VCHUER

retrovirus-related env polyprotein pseudogene - human

C. Species: Homo saptens (man)
C.Date: 28-Dec-1987 #sequence_revision 04-Jan-1996 #text_change 14-May-1999
C.Date: 28-Dec-1987 #sequence_revision 04-Jan-1996 #text_change 14-May-1999
C.Accession: E24483
J. Virol. 60, 589-598, 1986
A. Title: Nuclectide sequence of human endogenous retrovirus genome related to the mou A.Reference number: A93023; MUID:87036922
A.Accession: E24483
A.Acce

Gaps 1; Length 584; Indels Score 768.5; DB 4; Pred. No. 4e-66; 1; Mismatches 6; Query Match 78.6%; Best Local Similarity 94.7%; Matches 142; Conservative

ï

1 PVTWMDNPIEVYVNDSIWVPGPIDDRCPAKPEEEGMMINISIGYRYPPICLGRAPGCLMP 60 4 PVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMP 63 QQ à

64 AVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG 123 ò

q

124 SKNTEVLVWEECVANSVVILQNNEFGTIID 153 ò

121

probable transcription elongation factor [imported] - Arabidopsis thallana probable transcription elongation factor [imported] - Arabidopsis thallana (Mouse-ear cress)

C; Speckles: Arabidopsis thallana (Mouse-ear cress)

C; Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Mar-2001

C; Accession: T02309; G84753

R; Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K submitted to the EMBL Data Library, July 1998

A; Reference number: 214657

A; Reference number: 214657

A; Reference number: Capable (Fight Capable)

A; Reference number: Lansiated from GB/EMBL/DDBJ

A; Residues: 1-990 <ROU>

11;

Gaps

52;

Indels

Length 348;

```
transposase tnpA - Enterococcus faecalis plasmid pAD2 transposon Tn917
N.Alternate names: erythromycin resistance protein, ORF5; erythromycin resistance pro
C.Species: Enterococcus faecalis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R;Albright, D.; Balley, E.; Woodward, J.G.
Immunogenefices 34, 136-188, 1991
A;Title: Nucleotide sequence of a cDNA clone of the horse (Equus caballus) DRA gene.
A;Reference number: A53273; MUID:91331619
                                                                                                                                                                                                                                                                                                                                            A;Description: EC 2.7.7.6 [validated; MUID:20009333]
C;Superfamily: Saccharomyces cerevisiae DNA-directed RNA polymerase 40K chain
C;Keywords: nucleotidyltransferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C;Species: Equus caballus (domestic horse)
C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 21-Jan-2000
C;Accession: A53273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C;Superfamily: class II histocompatibility antigen; immunoglobulin homology F;97-162/Domain: immunoglobulin homology <IMM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EVPTV----SPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFR---PKGKTCPKEIPKG 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SKNTEV-----LVWEECVANSVVILQNNEFGTIIDLGTSRS-------ILPQLLR 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        272 KKQARVADVRKDTVSREC-----LRHPEFADKVQLGRVRDHYLFSVESTGIMKPDVLF 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EVYVNDSVWVP-GPTDDRCPAKPEEEGMMINISIGYHYPPICLGR-APGCLMPAVQNWLV 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MHC class II histocompatibility antigen DR alpha chain - horse (fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 VTPVTWMDN--PIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI----- 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Molecule type: mRNÅ
A;Residues: 1-226 <ALB>
A;Cross-references: GB:M60100; NID:g164236; PIDN:AAA30956.1; PID:g164237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YDCKVEHWGLDEPLLKHWEFEAPT--PLSETTENVVCGLGL 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53 -- CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 82; DB 2;
Pred. No. 1.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 2;
                                                                       A; Cross-references: EMBL: AF082512; PIDN: AAD44503.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 8.5%; Score 83; DB 3
Best Local Similarity 23.5%; Pred. No. 2.2;
Matches 46; Conservative 28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.4%; Score 82;
27.7%; Pred. No.
                                                                                                             A; Experimental source: strain 972h(-)
                                                                                                                                                            C;Genetics:
A;Gene: rpa42; SPBC1289.07c; rpc40+
A;Map Position: 2
C;Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           325 IKSIAVLKSKCLAVKS 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                167 TNSVV---SKCTSESS 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 28; Conserv
                            A; Residues: 1-348 <SHP>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: A53273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Map position: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A:Gene: DRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
A; Experimental source: cultivar Columbia

A; Experimental source: cultivar Columbia

R; Lin, X; Kaul, S; Kounsley, S.D; Shea, T.P; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 76.768, 1999

A; Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A; Reference number: A84420; MUID: 20083487

A; Reterence number: A84420; MUID: 20083487

A; Reterence number: Bay Signification

A; Residues: 1-990 <STO>
A; Residues: 1-990 <STO>
A; Residues: 1-990 <STO>
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Cross-references: GB: A26/3; 280/3; 308/3; 332/2; 451/3; 503/3; 536/2; 552/3; 603/3; 6
A; Introns: 97/1: 177/3; 226/3; 280/3; 308/3; 332/2; 451/3; 503/3; 536/2; 552/3; 603/3; 6
A; Introns: 97/1: Introns
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA-directed RNA polymerase (EC 2.7.7.6) chain Rpc40 [validated] - fission yeast (Schizch Ni Alternate names: rpa42 protein (Species: Schizosaccharonyces pombe (Species: Schizosaccharonyces pombe (Species: Schizosaccharonyces pombe (Species: Schizosaccharonyces) pombe (Species: O3-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000 (Species: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000 (Species: T43505; T50460 (Species: T43505; T50460 (Species: T43505; T50460 (Species: T43505; T50460 (Species: T3356 (Species: T3384 (Species: T3356 (Species: T336 (Sp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Reference number: 22552
A; Reference number: 22552
A; Reference number: 22552
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-348 < NGO;
A; Reference number: 225072
A; Reference number: 225072
A; Reference number: 225072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                108 KFRPKGKT-------CPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGT 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60 CLMPAVQNWLVEVP---TVSPNSRFTYHMVSGMSLRPRVNYL------QDFSYQRSL 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PMTPSSTSYLPTTPGGQAMTPGTDLDVMSLDIVNIFQFTDYVSLFFLCGHHQDGSCVVAL 917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      806 PGSWGTSPYEAATPGSDWGSSTPGRSSYRDAGTP-----INNGFVYYLLCLNANAPS 857
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 PVTWMDNPIEVYVNDSVW---VPGPTDDRCPAKPEEEGMMINISIGYHYPPICL-GRAPG 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 990,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8.5%; Score 83.5; DE
21.9%; Pred. No. 6.9;
tive 41; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SLDI -- KILKL-TILAKLVHE 990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SRSILPQLLRTNSVVSKCTSE 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

5.

Length 226

ŝ

Length 1003;

DB 1;

Indels

```
A;Status: translated from GB/EMBL/DDBJ
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-295 cR0U>
A;Cross-references: EMBL:AC002388; NID:g3420042; PID:g2344890
A;Cross-references: EMBL:AC002388; NID:g3420042; PID:g2344890
A;Experimental source: cultivar Columbia
R;Exin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujil, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujil, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shea, T.P.; Salzberg, S.L.; Fraser, C.M.; Venter
Misture 402, 761-7768, 1999
A;Tille: Sequence and analysis of Chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: R94884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           probable AP2 domain transcription factor [imported] - Arabidopsis thaliana N;Alternate names: transcription factor TINY homolog T13E15.5 (Species: Arabidopsis thaliana (mouse-ear cress) (C) Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 16-Feb-2001 (C) Accession: T00399; F84884 (S) Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, Submitted to the EMBL Data Library, July 1997 (A) Description: Arabidopsis thaliana chromosome II BAC T13E15 genomic sequence. A) Reference number: 214146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Cross-references: GB: AE002093; NID: 94895256; PIDN: AAD32841.1; GSPDB: GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hypothetical protein Y17G7B.10b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Species: T26502
C;Accession: T26502
R;Smye, R.
submitted to the EMBL Data Library, June 1998
                                                                                                                                                                                                               CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               68 WLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKT-CPKEIPKGSKN 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125 WLGTYPTAEMAAR--AHDVAALAIKGTTAYL---NPPKLAGELPRPVTNSPKDI-QAAAS 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           179 LAAVNWQD----SVNDVSNSEVAEIVEAEPSRAVVAQLFSSDTSTTTTTQSQEYSEASC 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    127 TEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSVVSKCT----SESSC 180
                                                                                                                                                                                                                                                                                                                                                                                        111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI
                                                                                                                                                                                                                                                                           441 CLKPKVPKNLRLWGWILDCDT----SRFIKHMADGSDDLDLDVRLN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Reference number: 220225
A;Accession: T26502
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.1%; Score 79; DB: 24.4%; Pred. No. 4.4; tive 29; Mismatches
                                                                         8.2%; Score 80.5; Ilarity 24.3%; Pred. No. 14; Conservative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 24.4%
Matches 29, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A,Gene: T13E15.5; At2g44940
A,Map position: 2
                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Residues: 1-295 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A, Molecule type: DNA
                                                                             Query Match
Best Local S:
Matches 27,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                             483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         T26502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A.C. Comment: This protein is a major component of the occlusion body which serves to prote C.Comment: This protein is a major component of the occlusion body which serves to prote C.Comment: This protein is a major component of the occlusion body which serves to prote C.S. Superfamily: Ansacta moreil povavirus spheroidin as parendin leucine zipper C.S. 21003/Product: spheroidin #status experimental <WAT>
F.119-140/Region: leucine zipper motif F.2-Modified site: acetylated amino end (Ser) (in mature form) #status experimental F.47.109.204.210,259.510.529,547.561.799,901,922.955/Binding site: carbohydrate (Asn) (covalent) #status absent
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 26-Aug-1999
C;Accession: JQ0971; E25028; F25028
R$An, FY: Clewell, D.B.
Plasmid 25, 121-124, 1991
A;Title: Tn917 transposase, sequence correction reveals a single open reading frame corn
A;Reference number: JQ0971; MUID:91312951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: protein
A; Residues: 2-7, 'X', 9-11, 'X', 13; 86-90, 'X', 92-98, 'X', 100-102; 532-540; 728-750; 786-802; 810-
B; Hall, R.L.; Moyer, R.W.
J. Virol. 65, 6516-6527, 1991
A; Title: Identification, cloning, and sequencing of a fragment of Amsacta moorei entomop
A; Reference number: A41561; MUID: 92046310
                                                                                                                                                                                                                                                                                                                                                                                                                                            macrolide-lincosamide-streptogramin B-resistand
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         from Amsacta moorei
                                                                                                                                                                                                                                A.Molecule type: DNA
A.Residues: 1-973 cANF>
A.Cross-references: GB:M11180; GB:M36722; NID:g154957; PIDN:AAA27455.1; PID:g154962
A.Experimental source: strain DS16
R:Shaw, J.H.; Clewell, D.B.
J. Bacteriol. 164, 782-796, 1985
A.FILLe: Complete nucleotide sequence of macrolide-lincosamide-streptogramin B-resis A.Reference number: A38108; MuID:86033641
A.Reference number: annotation; erratum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Species: Amsacta moorei poxylius
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
C; Accession: JQ1436; PQ0289; E41561
J; Gen. Virol. 73, 559-566, 1992
A; Title: The predicted amino acid sequence of the spheroidin protein from Am A; Reference number: JQ1436; MUID:92185464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      107 ILEVENYHRF------GSANTOKOLIEYLIE---LALENDDSIYLMKKTI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -LQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEF----GTI 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----- RPRVNY--- 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Molecule typė: DNA
Ksesidues: 1-1003 <BAN>
A;Cross.references: GB:M75889; NID:g209629; PIDN:AAA42378.1; PID:g209630
A;Accession: P00289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 973;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Note: this sequence has been revised in reference JQ0971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 2;
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 8.3%; Score 81; DB;
Best Local Similarity 21.8%; Pred. No. 12;
Matches 32; Conservative 26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     spheroidin precursor - Amsacta moorei poxvirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | | | | : ::|:| :: | 148 DFLTRKRIIFPSIATLEDIISRCRDKA 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                152 IDLGTSRSILPQLLRTNSVVSKCTSES 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Superfamily: transposase In21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Residues: 1-1003 <HAL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Accession: E41561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Genome: plasmid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Gene: tnpA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

9

Gaps

16;

Length 295; 45; Indels

DB 2;

Σ S.M.;

```
Query Match
Best Local Simi
Matches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical protein Y17G7B.10a - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Species: T5-oct-1999 #sequence_revision 15-oct-1999 #text_change 15-oct-1999
C.Accession: T26501
R.Smyc, R.
submitted to the EMBL Data Library, June 1998
A.Reference number: 220225
A.Accession: T26501
A.Accession: T2670
A.Acc
A; Residues: 1-492 <WIL>
A; Cross-references: EMBL:AL023828; PIDN:CAA19456.1; GSPDB:GN00020; CESP:Y17G7B.10b
A; Experimental source: clone Y17G7B
C; Genetics: CESP:Y17G7B.10b
A; Gene: CESP:Y17G7B.10b
A; Map position: 2
A; Introns: 12/3; 60/1; 124/3; 154/2; 240/2; 298/2; 423/1
                                                                                                                                                                                                                                                                                                                                                                    12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -VEIPEVEKETVELPVLALNEVFI----GESLSSRVSYYEIGINDAQMLKQKSSGIT-- 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             272 PATTRWTDNP-----PPVKKRMISEA- 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQ-DFSYQRSLKFRPKGKTCP 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KEIPKGSKNTEVLVW-----EECVANSVVILQNNEFGTIIDLGTSRSILPQLLR--T 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --ICTGTGSTS---WNFNINKLTEQCVQDLMKIV------AEHCNLPQIPHGDK 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     294 PATTRWTDNP------PSPAREIEECMSLS------PPVKKRMISEA- 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQ-DFSYQRSLKFRPKGKTCP 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KEIPKGSKNTEVLVW------EECVANSVVILQNNEFGTIIDLGTSRSILPQLLR--T 167
                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                              4 PVT--WMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI---CLGRAP 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 PVT--WMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI---CLGRAP 58
                                                                                                                                                                                                                                                                                                                                                                    :99
                                                                                                                                                                                                                                                                                                 1 8.1%; Score 79; DB 2; Length 492; Similarity 25.0%; Pred. No. 8.1; Conservative 23; Mismatches 52; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        / Match 8.1%; Score 79; DB 2; Length 514; Local Similarity 25.0%; Pred. No. 8.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          168 NSVVSKCT 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAVSEICT 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NSVVSKCT 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |:| ||
NAVSEICT 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  47;
                                                                                                                                                                                                                                                                                                                                                                    47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S:
Matches 47,
                                                                                                                                                                                                                                                                                                 Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 168
                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

10

RESULT

```
Rigorman, D.M.; Itoh, N.; Kitamura, T.; Schreurs, J.; Yonehara, S.; Yahara, I.; Arai, Proc. Natl. Acad. Sci. U.S.A. 87, 5459-5463, 1990
A/Title: Cloning and expression of a gene encoding an interleukin 3 receptor-like pro A; Reference number: A35782; MUID:90319131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: mRNA
A; Residues: 1-896 < GORS
A; Residues: 1-896 < GORS
A; Cross-references: GB:M34397; NID:9191821; PIDN:AAA37204.1; PID:9309101
A; Cross-references: GB:M34397; NID:9191821; PIDN:AAA37204.1; PID:9309101
C; Comment: Mouse high-affinity IL-5, GM-CSF, and one class of high-affinity IL-3 recef.
C; Superfamily: interleukin-3 receptor beta chain: cytokine receptor homology
C; Keywords: cytokine receptor; duplication; transmembrane protein
F; 12-2/Domain: signal sequence #status predicted < SIG>
F; 23-866/Product: cytokine receptor common beta chain #status predicted < MAT>
F; 33-441/Domain: cytokine receptor homology < CRSI>
   E O
                                                                                  Cyaccession: T50303
RyMcDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Saunders, D.; Harris, D.
submitted to the EMBL Data Library, January 2000
A; Reference number: 225059
A; Accession: T50303
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-704
A; Cross-references: EMBL:ALl36521; PIDN:CAB66312.1; GSPDB:GN00066; SPDB:SPAPYUG7.03c
A; Experimental source: strain 972h(-); clone plasmid pYUG7
hypothetical protein SPAPYUG7.03c [imported] - fission yeast (Schizosaccharomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11;
                                  C;Species; Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
C;Accession: A35782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 70 VEVPTVSPNS---RFTYHMVSGMS----LRPRVNYLQDFSYQ-----RSLKFRPK---- 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  268 DENRENCRIDIYRSDSIHE---YEEPLISSITSLDSPH---VLDENAPIPLLPKVVSLPD 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10 NPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNWL 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VPG-PTDDRCPAKPEEEGMMINISIG-YHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNS 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 704;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F:442-463/Domain: transmembrane #status predicted <TWM>F:464-896/Domain: intracellular #status predicted <INT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cytokine receptor common beta chain precursor - mouse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;253-434/Domain: cytokine receptor homology <CRS2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22.3%; Pred. No. 24;
tive 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       322 PRFTNVLSAFDALTRTYLLRQNSKVVHATSQ 352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.8%; Score 76;
Similarity 25.9%; Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 7.8%
Best Local Similarity 22.3%
Matches 47; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Gene: SPDB:SPAPYUG7.03c
A;Map position: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PP
```

S

```
Cyclooxygenase 1 - rat
Cyclooxygenase 1 - rat
Cyclooxygenase 1 - rat
Cyclooxygenase 1 - rat
Cyclooxygenase 10-rat
Cyclooxygenase 10-rat
Cyclooxygenase 10-bec-1999
Cyclooxygenase 10-be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C.D.; Cho, T.J.; Thanh, V.H.; Scallon, B.J.; Fischer, R.L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ς.
.`
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Glycinin G3 precursor - soybean G) Species: Glycine max (soybean) G) Species: Glycine max (soybean) G) Species: Glycine max (soybean) G; Date: 21-Nov-1993 #sequence_revision 19-Jan-1996 #text_change 03-May-1996 G; Accession: S1003 R; Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallon, B.J.; Fiss Plant Cell 1, 313-328, 1989 Plant Cell 1, 313-328, 1989 Plant Cell 1, 313-328, 1985 
                                                                                                                                                                                                                           :::||:|:
83 SYTNAPQEIYIQQGSGIFGMIFPGCPSTFEEPQQKGQSSRPQDRHQXIYHFREGDLIAVP 142
                                                                                                                                                                        102 SYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSIL 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 45 IGYHYPPICLGRAPGCLMPAVQNWLV---EVPTVSPNSRFTYHWVSGMSLRPRVNYLQDF 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   143 TGFAY---------WMYNNEDTPVVAVSLIDTNSFONOLDOMPRRFYLAGN 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          102 SYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSIL 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            185 QEQEFLQYQPQKQQGGTQSQKGKRQQE------EENEGGSILS----GFA 224
45 IGYHYPPICLGRAPGCLMPAVONWLV---EVPTVSPNSRFTYHMVSGMSLRPRVNYLQDF 101
                                                                                  -----WMYNNEDTPVVAVSLIDTNSFQNQLDQMPRRFYLAGN 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----EGMMINIS 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Score 75; DB 2; Length 484;
; Pred. No. 19;
28; Mismatches 67; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Keywords: storage protein
F;1-19/Domain: signal sequence #status predicted <SIG>F;20-484/Product: glycinin G3 #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6 TWMDNPIEVYVNDSVWVPGPTDDRCPAKPEE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.7%; Score 75;
18.2%; Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 18.2
15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1: | | | 3
225 PEFLEHAFVVDR 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        162 PQLLRTNSVVSK 173
                                                                                                                                                                                                                                                                                                                                                 162 PQLLRTNSVVSK 173
                                                                                                                                                                                                                                                                                                                                                                                                                                   225 PEFLEHAFVVDR 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Superfamily: glycinin
                                                                                      143 TGFAY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Matches 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                               q
                                                                                                                                                                        ò
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Residues: 1-318 <KAN>
A; Residues: 1-318 <KAN>
A; Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAA10147.1; PID:d101079
A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Species: Glycine max (soybean)
C; Species: Glycine max (soybean)
C; Species: Glycine max (soybean)
C; Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Aug-1999
C; Accession: S04665
A; Title: The glycinin Gy(3) gene from soybean.
A; Reference number: S04665; MUID: 89296500
A; Residues: 19481 < CHO>
A; Residues: 1-481 < CHO>
A; Cross_references: EMBL: X15123; NID: 918638; PIDN: CAA33217.1; PID: 918639
A; Experimental source: variety Dare
C; Genetics:
                                                                                                                                                                                                                                                                                                                                        hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             207 MQGRIDAFAGDGILLFG------EALVLDIPLGRSYRLYPPYPLDCQGY--GLI 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S SYTNAPQEIYIQQGSGIFGMIFPGCPSTFEEPQQKGQSSRPQDRHQKIYHFREGDLIAVP 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62 MPAVQ-NW--LVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPK 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIG--YH-YPPI----CLGRAPGCL 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51; Indels 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62;
                                                                                               846 E-TEDLVQDLSVKKFPYQPMPQAPAIQFFKSLKHQDYLSLPPWDNSQSGKVC 896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2; Length 318;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 481;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67;
                                             -----PRVNYLQDFSYQRSLKFRP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2;
19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 TWMDNPIEVYVNDSVWVPGPTDDRCPAKPEE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.7%; Score 75.5; DE 27.6%; Pred. No. 10; ative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.7%; Score 75; DB
18.2%; Pred. No. 19;
ative 28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: S74322; MUID: 97061201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 27.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A,Gene: Gy3
A,Introns: 96/1; 177/3; 352/3
C,Superfamily: glycinin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Status: preliminary
Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        119 EIPKGSK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        311 ESPSGKE 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C; Accession: S76295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: S76295
                                         80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

```
8;
                                                  Query Match 7.6%; Score·74.5; DB 2; Length 602; Best Local Similarity 23.3%; Pred. No. 28; Matches 37; Conservative 16; Mismatches 61; Indels 45; Gaps
                        Oy.
                                                   ò
                                    g
```

Search completed: April 9, 2002, 16:58:15 Job time: 249 sec

Appli Appli Appli Appli Appli Appli Appli Appl1

```
10, Appl
10, Appl
2, Appli
2, Appli
2, Appli
2, Appli
2, Appli
451, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 6, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Mayer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System NUMBER OF SEQUENCES: 66 CORRESPONDENCE ADDRESS:
APPLICANT: SEQUENCES: 66 CORRESPONDENCE ADDRESS:
ADDRESSED: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
                      Sequence 2
Sequence 5
Sequence 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 3266
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
TASSIFICATION: 435
                                    5180868-2

US-08-726-525-2

US-08-726-0368-2

US-09-038-516-2

US-09-268-754-451

US-08-261-304-2

US-08-261-304-2

US-08-261-304-2

US-08-261-304-2

US-08-317-5228-5

US-08-738-975-5

US-08-738-975-5

US-08-738-975-5

US-08-738-626-5

US-08-08-626-5

US-08-08-626-5

US-08-08-626-5

US-08-08-626-5

US-08-08-626-5
US-08-313-288B-10
PCT-US93-03164-10
                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: SAliwanchik, David R.
RESISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C3
TELECOMUNICATION INCORMATION:
TELECHHONE: 904-375-8100
TELEFAX: 904-375-800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1003 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-991-8678-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
      RESULT 1
US-07-991-867B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH:
  (without alignments)
117.279 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ), App
Appl
Appl
Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 6, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 100,
Sequence 100,
                                                                                                                                                                                                                                                  1 MVTPVTWMDNPIEVYVNDSV......PQLLRTNSVVSKCTSESSCR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 100
Sequence 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 6,
Sequence 6,
Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 14
Sequence 14
Sequence 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1
Sequence 1
Sequence 2
Sequence 3
Sequence 3
                                                                                                                                 April 9, 2002, 16:57:28; Search time 34.73 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                    GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-460-998-12
US-08-086-428B-100
US-08-290-655A-100
PCT-US95-10398-100
US-09-577-332
US-08-857-076-99
US-08-857-076-99
US-08-131-118-15
US-08-413-118-15
US-08-473-446-15
US-08-290-731C-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-896-449A-2

US-08-132-652-2

US-08-218-265-14

US-08-218-265-14

US-08-521-872-14

US-08-590-399-14

US-07-930-686-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-700-013B-27
US-07-862-021B-10
                                                                                                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                   212252 segs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                              OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                           US-09-490-700-40
978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2813
1420
1420
1420
1420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73
69
69
68
68
68
67
67
                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
No.
```

```
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2
                                                                                                                                                                                               60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruid! Michael E.
TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29;
8.2%; Score 80.5; DB 1; Length 1003; 24.3%; Pred. No. 1.2; Live 16; Mismatches 39; Indels 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 8.2%; Score 80.5; DB 1; Length 1003; Best Local Similarity 24.3%; Pred. No. 1.2; Matches 27; Conservative 16; Mismatches 39; Indels 29
                                                                                                                                                                                                                                                                                                                                                                                         111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    483 -RNDICLKQAIKQHYTNVIILEYANTYPNCTLSLGNNRFNNVFDMNDNKTI 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                     : | |: | : | : | : | 483 -RNDICLKQAIKQHYTNVIILEYANTYPNCTLSLGNNRFNNVFDMNDNKTI 532
                                                                                                                                                                                                                                                                          441 CLKPKVPKNLRLWGWILDCDT---SRFIKHMADGSDDLDLDVRLN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         441 CLKPKVPKNLRLWGWILDCDT---SRFIKHMADGSDDLDLDVRLN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
THE ADAR TO THE A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Saliwanchik, David R. REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 6, Application US/08107755A Patent No. 5721352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1003 amino acids
                                             Best Local Similarity 24.3%
Matches 27; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-107-755A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gainesville
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE: Florida
COUNTRY: U.S.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
US-08-107-755A-6
Query Match
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
```

```
60 CLMPAVQ-----NWLVEVPTVSPNSRFTYHMVSG---MSLRPRVNYLQDFSYQRSLKFR 110
                         APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H Procession System
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1003;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      111 PKGKTCPKE-IPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          441 CLKPKVPKNLRLWGWILDCDT --- SRFIKHMADGSDDLDLDVRLN ---
                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Score 80.5; DB; pred. No. 1.2; 16; Mismatches
                                                                                                                                                                                                                  ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-MG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: W0 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-7AM-1992
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNEL/CALL ALIGHER 185,746
REGISTRATION NUMBER: 35,746
REFRENCE/SDOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-D6C-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 07/657,584 FILING DATE: 19-FEB-1991 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC_DOS/MS-DOS
US-08-544-332-6; Sequence 6, Application US/08544332; Patent No. 5935777; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 8.2%;
Best Local Similarity 24.3%;
Matches 27; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 1003 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      904-372-5800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-08-544-332-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE:
                                                                                                                                                                                                                                                                                                                                    32606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
```

5,

RESULT

.. 8

```
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS TITLE OF INVENTION: OF USE CURRENT APPLICATION OF USE CURRENT APPLICATION NUMBER: US/09/132,652 CURRENT FILING DATE: 1998-08-11 EARLIER FILING DATE: 1998-08-11 EARLIER FILING DATE: 1997-07-18 NUMBER OF SEQ ID NOS: 29 SOFTWARE: PATENTI VET: 2.0 SEQ ID NO 2 LENGTH: 2813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        75 VSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVL---- 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          131 -- VWEECVANSVVI-------LONNEFGTI------IDLGTSRSILPQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .859 TDHVCDATCSAIGMAHYLTFDGLKYLFPGECQYVLVQDYCGSNPGTLRILVGNEGCSYPS 918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26 TDDRCPAKPEEEGMMINIS---IGYHYPPIC-----LGRAPGCLMPAVQNWLVEVPT 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 14, Application US/08540804

Patent No. 591966
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Koleske, Anthony J.
APPLICANT: Chao, David M.
TITLE OF INVENTION: No. 5919666el Factors Which Modify Gene TITLE OF INVENTION: Transcription and Methods of Use Therefor NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C. STREET: Two Militia Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWAREN PAINTED THE CONTROL #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540,804
FILING DATE: 11-OCT-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08 08/521,872
FILING APPLICATION DATA:
APPLICATION NUMBER: US 08/218,265
FILING DATE: 25-MAR-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 8.2%; Score 80.5; DB 3; Best Local Similarity 21.8%; Pred. No. 5.5; Matches 46; Conservative 29; Mismatches 73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1031 CADTKKVPLDSSPAVCHNNIMKQTMVDSSCR 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       164 LLRT------NSVVSKCTSESSCR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LIF: VALIA
COMPUTER READABLE FORM:
MEDIUW TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
     Duffendeck, John
                                                                                                                                                                                                                                                                                        ; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-132-652-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CITY: Lexington
STATE: Massachusetts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 02173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-540-804-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  859 TDHVCDATCSAIGMAHYLTFDGLKYLFPGECQYVLVQDYCGSNPGTLRILVGNEGCSYPS 918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       75 VSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVL---- 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     919 VKCKKKVTI-LVEG------GEIELFDGEVNVKKPMKDETHFEVVESGQYVILLLGKAL 970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      131 --VWEECVANSVVI------LQNNEFGTI------IDLGTSRSILPQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26 TDDRCPAKPEEEGMMINIS---IGYHYPPIC-----LGRAPGCLMPAVQNWLVEVPT 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match

8.2%; Score 80.5; DB 3; Length 2813;
Best Local Similarity 21.8%; Pred. No. 5.5;
Matches 46; Conservative 29; Mismatches 73; Indels 63
                                                                                          APPLICANT: Venta, Patrick J
APPLICANT: Yuzbasiyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
TITLE OF INVENTION: FACTOR AND METHODS OF USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,449A
                                                                                                                                                                                                                                                                                 ADDRESSEE: Harness, Dickey & Pierce, P.L.C. STREET: 5445 Corporate Drive CITY: Troy ACATE: Mchigan COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1031 CADTKKVPLDSSPAVCHNNIMKQTMVDSSCR 1061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 248-641-1600
TELEFAX: 248-641-0270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NS-09-132-652-2
Sequence 2, Application US/09132652
Fatent No. 6074832
GRNERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: YLEDSASiyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                 Sequence 2, Application US/08896449A Patent No. 6040143 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 287637

TELEX: 287637

INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 2813 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 18-JUL-1997
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATTORNEY/AGENT INFORMATION:
NAME: Smith, DeAnn F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein
                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      164 LLRT-----
                                                                                                                                                                                                                                                                                                                                                                                                       48098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-896-449A-2
US-08-896-449A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

amino acid

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      69 INDDPVPTPPAIEHKPVSPDKIGTFTADYSKPNLPPHYALFLKALRRKIYINLALGSHNK 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | | | | | : | | :| | : | : | : | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 PICLGRAPGCL-MPAVQNWLVEV-PTVSPNSRFTYHM-VSGMSLRP--RVNYLQDFSYQR 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106 SLKFRPKG-----KTCPKE--IPKGSKNTEVLVWEECVANSVVILQNNE----FGTIIDL 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Young, Richard A.
APPLICANT: Thompson, Craig M.
TITLE OF INVENTION: No. 5922585el Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Thereof
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1420;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READBABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 7.5%; Score 73.5; I
Best Local Similarity 23.0%; Pred. No. 14;
Matches 47; Conservative 29; Mismatches
                                                                                    WHI94-03A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH194-03
TELECOMUNICATION INFORMATION:
TELEPHONE: 617-861-5240
TELEFAX: 617-861-5540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US/08/218,265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 14, Application US/08218265 Patent No. 5922585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              246 GHLNGHTPTIASYLTPLLEAKKLV 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             155 G-----TSRSILPQLLRTNSVV 171
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFRENCE/CDCKET NUMBER: WHIS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 14: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1420 amino acids
                                                                                                                                                                                                                                                                                                                              1420 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: protein US-08-540-804-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
FILING DATE: 25-MAR
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ns
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               02173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 7
US-08-218-265-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    69 INDDPVPTPPAIEHKPVSPDKIGTFTADYSKPNLPPHYALFLKALRRKIYINLALGSHNK 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            51 PICLGRAPGCL-MPAVQNWLVEV-PTVSPNSRFTYHM-VSGMSLRP--RVNYLQDFSYQR 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              106 SLKFRPKG-----KTCPKE--IPKGSKNTEVLVWEECVANSVVILQNNE----FGTIIDL 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77; Indels 51; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                  ---MMINISIGYHYP 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: Young, Richard A.
APPLICANT: Young, Anthony J.
APPLICANT: Chao, David M.
APPLICANT: Chao, Carig M.
APPLICANT: Chao, David M.
APPLICANT: Chao, Carig M.
APPLICANT: Chao, Chao
                                                                                                                                                                                                                   Length 1420;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E: Hamilton, Brook, Smith & Reynolds, P.C.
Two Militia Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: PATENTIN FELGASE #1.0, Version #1.30 CURRENT APPLICATION DATA:
FILING DATE: 31-AUG-1995
CLASSIFICATION: 436
CLASSIFICATION: 436
FILING DATE: 35-MAR-1994
ATTORIEX_AGENT INFORMATION:
APPLICATION NUMBER: US 08/218,265
FILING DATE: 25-MAR-1994
ATTORIEX_AGENT INFORMATION:
                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                       Score 73.5; I
Pred. No. 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         246 GHLNGHTPTIASYLTPLLEAKKLV 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Sequence 14, Application US/08521872
; Patent No. 6015682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    155 G-----TSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                          29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                             16 VNDSVWVPGPTDDRCPAKPEEEG-
                                                                                                                                                                                                                   Query Match
Best Local Similarity 23.0%;
Matches 47; Conservative 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1420 amino acids
; TOPOLOGY: linear; MOLECULE TYPE: protein US-08-218-265-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE: protein US-08-521-872-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Massachusetts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: Two Mili
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            02173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 8
US-08-521-872-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
```

11;

```
51 PICLGRAPGCL-MPAVQNWLVEV-PTVSPNSRFTYHM-VSGMSLRP--RVNYLQDFSYQR 105
                                                                                                                                                                                                                                                                                                                 106 SLKFRPKG-----KTCPKE--IPKGSKNTEVLVWEECVANSVVILQNNE----FGTIIDL 154
                                                                                                                         ---MMINISIGYHYP 50
                             DB 4; Length 1420;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 69.5; DB 1; Length 440; Pred. No. 7.4;
                                                                        77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/930,686
FILING DATE: 19921006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INCORNATION:
APPLICANT: Sharp, Phillip J
APPLICANT: Cobon, Gary S
TITLE OF INVENTION: Nematode Vaccine
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS: 3
ADDRESSEE: Foley and Lardner
STREET: Suite 500, 1800 Diagonal Road
CITY: Alexandria
COUNTRY: United States of America
ZIP: 22213-0299
COMPUTER READABLE FORM:
MEDIUM TIPE: IDPPPY disk
COMPUTER: IBM PC COMPALIBLE
COMPATING SYSTEM: PC-DOS/MS-DOS
                        Query Match 7.5%; Score 73.5; D
Best Local Similarity 23.0%; Pred. No. 14;
Matches 47; Conservative 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 19921006
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK4486
FILING DATE: 06-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00040
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 12, Application US/07930686 Patent No. 5525508
                                                                                                                                                                                                                                                                                                                                                                                                           155 G-----TSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                        246 GHLNGHTPTIASYLTPLLEAKKLV 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELECOMMUNICATION INFORMATION TELEPHONE: (703) 836-9300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12:
                                                                                                                         16 VNDSVWVPGPTDDRCPAKPEEEG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.18; 22.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          : 440 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (703) 683-4109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: 899149
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Bent, Stephen A REGISTRATION NUMBER: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE:
US-07-930-686-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-07-930-686-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX:
                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                 11;
                                                                                                                                                                                                                   69 INDDPVPTPPAIEHKPVSPDKIGTFTADYSKPNLPPHYALFLKALRRKIYINLALGSHNK 128
                                                                                                                                                                                            51 PICLGRAPGCL-MPAVQNWLVEV-PTVSPNSRFTYHM-VSGMSLRP--RVNYLQDFSYQR 105
                                                                                                                                                                                                                                                                                      106 SLKFRPKG-----KTCPKE--IPKGSKNTEVLVWEECVANSVVILQNNE----FGTIIDL 154
                                                                                                -----MMINISIGYHYP 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Young, Richard A.
APPLICANT: Young, Richard A.
APPLICANT: Thompson, Craig M.
APPLICANT: Thompson, Craig M.
APPLICANT: Thompson, Craig M.
APPLICANT: Chao, David M.
TITLE OF INVENTION: No. 6214588el Factors Which Modify Gene
TITLE OF INVENTION: Transcription and Methods of Use Therefor
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STREET: Massachusetts
COUNTRY: USA
                                                 51;
Length 1420;
                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/590,399
FILING DATE: 26-JAN-1996
CLASSIFICATION ADATA:
APPLICATION NUMBER: US 08/540,804
FILING DATE: 11-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/521,872
FILING DATE: 31-AUG-1995
PRIOR APPLICATION NUMBER: US 08/521,872
FILING DATE: 31-AUG-1995
FILING DATE: 31-AUG-1995
FILING DATE: 25-MAR-1994
ATPONDRY/AGERT INFORMATION:
AMANE: CARRANATION DATE: 25-MAR-1994
ATPONDRY/AGERT INFORMATION:
    DB 3;
Query Match 7.5%; Score 73.5; Di
Best Local Similarity 23.0%; Pred. No. 14;
Matches 47; Conservative 29; Mismatches
                                                                                              16 VNDSVWVPGPTDDRCPAKPEEG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER: 32,227
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH194-03A3
TELECHMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 14:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 14, Application US/08590399 Patent No. 6214588 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                    155 G-----TSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                 246 GHLNGHTPTIASYLTPLLEAKKLV 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1420 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; MOLECULE TYPE: protein US-08-590-399-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-590-399-14
                                                                                                                                                                                               ç
                                                                                                                                                                                                                                         පි
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

9

```
108 --KFRPKGKTCPKEIPKGSKNTEV---LVW-EECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                          64 AVQNWLVEVPTVSPNSRFTYH-----MVSGMSLRPRVNYLQDFSYQRSL----- 107
                                                                                                                                                                                                                                    221 PT----TEKPTTTPK---IYHPGGMCPENNNGMTDEARQMFVDKHNEYRSLIAKGLAHNN 273
                                                                                                                                                                                                                                                                                                                                                                                               274 LGGFAPKA----ARMMKVSYNCEIEANRVEWAKDCTLGYNSVAQNNQWGYNV----HSL 324
    10 NPIEVYVNDSVWVPGPTDDRCP-----AKPEEEGMMINISIGYHYPPICLGRAPGCLMP 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26 TODRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQN-----WLVEVPTVS---- 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13 TND-CP-----NSSIVYEADNLIL-HAPGCVPCVRQNNVSRCWVQITPTLSAPNL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R.H.
NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
OF 51 ISOLATES OF HEPATITIS C AND THE USE
OF RAGENTS DERIVED FROM THESE SEQUENCES IN
DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 100, Application US/08086428B
Fatent No. 5514539
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION:
TITLE OF SEQUENCES:
TITLE OF SEQUENCES:
TITLE OF SEQUENCES:
NUMBER OF SEQUENCES:
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
CORPUSED ADDRESS ADDRESS:
CORPUSED ADDRESS:
CORPUSED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 7.0%; Score 68.5; Best Local Similarity 24.3%; Pred. No. 2.9 Matches 41; Conservative 19; Mismatches
                                                         177 NRINHYIYD-IGDPCTTDEDCQCTGCTCSKDE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | NAME: RICHARD W. BORK | REGISTRATION NUMBER: 36,459 | REFERNICE/DOCKET NUMBER: 2026-4070 | TELECOMMUNICATION INFORMATION: TELEPHONE: (212) 758-4800 | TELERAX: (212) 751-6849 | TELERAX: (212) 751-6849 | TELERAX: 421792 | TELERA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/08/086,428B FILING DATE: 29-JUN-1993 CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,4:
FILING DATE: 29-JUN-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E: FLOPPY DISK
IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           192 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: homosapiens
INDIVIDUAL ISOLATE: SA7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
STRANDEDNESS: unb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             || : :|
325 LPHINKT 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161 LPQLLRT 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-086-428B-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-086-428B-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET:
                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                          ò
                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
    11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11;
                                                                                                                          | || :| || 221 PT----TEKPTTTPK---IYHPGGMCPENNNGMTDEARQMFVDKHNEYRSLIAKGLAHNN 273
                                                                                                                                                                                                                                                                                                                                                                                               --KFRPKGKTCPKEIPKGSKNTEV----LVW-EECVANSVVILQNNEFGTIIDLGTSRSI 160
                                                                                                                                                                                                                                                                                                                                                                                                                                           274 LGGFAPKA----ARMMKVSYNCEIEANRVEWAKDCTLGYNSVAQNNQWGYNV----HSL 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                             10 NPIEVYVNDSVWVPGPTDDRCP-----AKPEEEGMMINISIGYHYPPICLGRAPGCLMP 63
    61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61;
                                                                                                                                                                                                                                64 AVQNWLVEVPTVSPNSRFTYH-----MVSGMSLRPRVNYLQDFSYQRSL-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.1%; Score 69.5; DB 2; Length 440; 22.5%; Pred. No. 7.4;
    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              59; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUPTY: United States of America
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
59;
    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Sharp, Phillip J
APPLICANT: Wagland, Barry M
APPLICANT: Cobon, Gary S
TITLE OF INVENTION: Nematode Vaccine
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSE: Foley and Lardner
STREET: suite 500, 3000 K Street, NW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA.
APPLICATION NUMBER: US 07/930,686
FILING DATE: 06-0CT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK4486
FILING DATE: 06-FEB-1991
PRIOR APPLICATION NUMBER: PCT/AU92/00040
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/460,998 FILING DATE: 05-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 12, Application US/08460998
Patent No. 5942413
GENERAL INFORMATION:
25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29,768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION TELEPHONE: (202) 672-5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEX: 904136,
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (202) 672-5300
(202) 672-5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           440 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Bent, Stephen A REGISTRATION NUMBER: 2
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein US-08-460-998-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      || : :|
LPHINKT 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161 LPQLLRT 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-460-998-12
42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Matches
                                                                                                                                                                                                                                                                                                                                                                                           108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

6

```
61 GAVTAPLRRANDYLAGGAALCSAL-YVGDACGAVFLVGQMFSYRPRQHTTVQDCNCSIYS 119
61 GAVTAPLRRAVDYLAGGAALCSAL-YVGDACGAVFLVGQMFSYRPRQHTTVQDCNCSIYS 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        77 ----PNSRFTYHMVSGMSLRPRVNYLQD-----FSYQRSLKFRPKGKT----CPKEIPK 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26 TDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQN-----WLVEVPTVS---- 76
                                                                                                                                                                                                                                                          N.H.

NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE I AND
CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
AND THE USE OF REAGENTS DERIVED FROM THESE
SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 192;
                                                                           ---PTTALVMAQLLRIPQVV 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123 GSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 GHITGHRMAWDMMNWS-------PITALVMAQLLRIPQVV 153
                                     123 GSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.0%; Score 68.5; DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.
                                                                                                                                                                                                                                           MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/08/290,665A FILING DATE: 15-AUG-1994 CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2026-4116
                                                                                                                                                                            Sequence 100, Application US/08290665A Patent No. 5882852 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/CDCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE: MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36,459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                       345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SA7
US-08-290-665A-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                          TITLE OF INVENTION: NUC.
TITLE OF INVENTION: AMIL
TITLE OF INVENTION: COM
TITLE OF INVENTION: SEQUENCES: 26
CORRESPONDENCE SEGUENCES: 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                             120 GHITGHRMAWDMMNWS--
                                                                                                                                                                                                                                           BUKH, J.,
PURCELL, R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: NEW YORK STATE: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                            US-08-290-665A-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10154
                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
g
                                                                        쉄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6
                ----PNSRFTYHMVSGMSLRPRVNYLQD-----FSYQRSLKFRPKGKT----CPKEIPK 122
                                                    61 GAVTAPLRRAVDYLAGGAALCSAL-YVGDACGAVFLVGQMFSYRPRQHTTVQDCNCSIYS 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----PNSRFTYHMVSGMSLRPRVNYLQD-----FSYQRSLKFRPKGKT----CPKEIPK 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26 TDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQN-----WLVEVPTVS---- 76
                                                                                                                                                                                                                                                                                                                                   NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
OF SI ISOLATES OF HEPATITIS C AND THE USE
OF REAGENTS DERIVED FROM THESE SEQUENCES IN
DIAGNOSTIC METHODS AND VACCINES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7.0%; Score 68.5; DB 2; Length 192; 24.3%; Pred. No. 2.9;
                                                                                                                             120 GHITGHRMAWDMMMWS------PTTALVMAQLLRIPQVV 153
                                                                                           123 GSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLERTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2026-4070US1
                                                                                                                                                                                                                                                                                                 MILLER, R.H. AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/468,570 FILING DATE: 6-JUN-1995 CLASSIFICATION: 424 PRIOR APPLICATION NUMBER: 08/086,428 FILING DATE: 29-JUN-1993 CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                   Sequence 100, Application US/08468570 Patent No. 5871962 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --
--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGIESTRATION NUMBER: 36,459
REFRENCE/COCKET NUMBER: 2026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MORGAN & FINNEGAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX: 421792
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    individual isolate: SA7 US-08-468-570-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               345 PARK AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: MUCITILE OF INVENTION: OF TITLE OF INVENTION: OF TITLE OF INVENTION: OF FOR INVENTION: DIAMOMBER OF SEQUENCES: 155 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
                                                                                                                                                                                                                                                                                             APPLICANT: BUKH, J., APPLICANT: PURCELL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
tes 41; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NEW YORK
: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                    JS-08-468-570-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11
                ð
                                                      g
                                                                                             ò
                                                                                                                                 염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
```

6

ω

```
9;
                                                                                     APPLICANT: BUCK, J., MILLER, R.H. AND
APPLICANT: BUCK, J., MILLER, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: CORE GENES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF TECOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 263
CORRESPONDENCE ADDRESS:
ADDRESSE: MORGAN & FINNEGAN
STRET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 GAVTAPLRRAVDYLAGGAALCSAL-YVGDACGAVFLVGQMFSYRPRQHTTVQDCNCSIYS 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               77 ----PNSRFTYHMVSGMSLRPRVNYLQD-----FSYQRSLKFRPKGKT----CPKEIPK 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26 TDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQN-----WLVEVPTVS---- 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13 TND-CP-----NSSIVYEADNLIL-HAPGCVPCVRQNNVSRCWVQITPTLSAPNL 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       58; Indels 51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 7.0%; Score 68.5; DB 5; Length 192; Best Local Similarity 24.3%; Pred. No. 2.9; Matches 41; Conservative 19; Mismatches 58; Indels 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123 GSKNTEVLVWEECVANSVVILQNNEFGTIIDLGTSRSILPQLLRTNSVV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29 JUNE 1993
PRIOR APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
APPLICATION NUMBER: 08/290/665
FILING DATE: 15 AUGUST 1994
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 2026-4116
TELECOMMUNICATION INFORMATION:
TELEFAN: (212) 758-4800
TELEFA: (212) 751-6849
TELER: (212) 751-6849
TELER: (212) 751-6849
TELER: (212) 751-6849
TELER: 421792
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 10154
COMPUTER REAABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: 1BM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10398
FILING DATE: 15-AUG-1995
CLASSIFICATION:
RESULT 15
PCT-US95-10398-100
; Sequence 100, Application PC/TUS9510398
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: homosapiens ; INDIVIDUAL ISOLATE: SA7 PCT-US95-10398-100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

Search completed: April 9, 2002, 16:57:29

Job time: 248 sec

```
16, Appl
3, Appl
3, Appl
3, Appl
5, Appl
10, Appl
10, Appl
10, Appl
10, Appl
39, Appl
12, Appl
13, Appl
14, Appl
16, Appl
17, Appl
18, Appl
18, Appl
19, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 523 EADSYLRA----QAGNQAQEEQMCEICLRGEDGPNMLLCDECNRGYHMYCLQ----PAL 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KKLQSFYLWEWEEKGIST-----PRPKIISPVSGPE------HPELWRLTVASHHIRI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DXAPRGQFYHNCSGQTQSCPSAQVSP----AVDSDLTESL-------DXHKH 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QDFSYQRSLKFRPKGKTCPKE----IPKGSKNTEVLVWEECVAN-SVVILQNNEFGTII 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
  Sequence 1
Sequence 5
Sequence 1
Patent No.
                                                                                                           Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
US-09-051-019-2
; Sequence 2, Application US/09051019
; Patent No. 6103229
; GENERAL INFORMATION:
; APPLICANT: KAHMANN, Regine and QUADBECK-SEEGER, Claudia
; TITLE OF INVENTION: Regulatory gene from Ustilago maydis
; WUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauf
STREET: 1101 Connecticut Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 2289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3.6%; Score 108.5; DB 3; Length; 17.8%; Pred. No. 0.13; 1ve. 77; Mismatches 195; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---NSRFTYHMVSGMSLRPRVN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mb storage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage COMPUTER: IBM AT-compatible, Pentium processor OPERATING SYSTEM: Windows 98 SOFTWARE: WordPerfect version 6.1 CURRENT APPLICATION DATA: APPLICATION DATA: APPLICATION NUMBER: US/09/051,019 FILLING DATE: 31-MAR-1998
               US-08-928-692-58
US-08-162-402B-16
5310678-1
US-08-937-236-3
US-08-937-236-2
US-08-937-236-2
US-08-937-236-2
US-08-937-236-2
US-08-937-236-2
US-08-133-288B-10
US-08-133-03164-10
US-08-670-707A-39
US-09-170-996-12
US-08-670-707A-37
US-09-170-996-12
US-08-670-707A-37
US-09-170-996-12
US-08-670-707A-37
US-09-170-996-13
US-08-670-707A-37
US-09-170-996-13
                                                                                                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2289 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 3.6
Best Local Similarity 17.8
Matches 118; Conservative
  863
916
159
1253
1291
1291
1291
1443
1443
11443
11876
2133
2133
793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              = ::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LVEVPTVSP----
  *************************
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
TOPOLOGY:
US-09-051-019-2
153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194
  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                          (without alignments)
363.500 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appli
, Appl
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               No. is the number of results predicted by chance to have a greater than or equal to the score of the result being printed, derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 5, 1
Sequence 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 14
Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2
Sequence 2
Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence (Sequence (Sequence (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence Sequence
                                                                                                                                                                                                                                                                                 .....VCRCTPTAPKKTVTSRTGHE
                                                                                                                                                   ; Search time 34.73 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                      GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT - US93 - 03164 - 14
US - 08 - 202 - 841A - 2
US - 08 - 202 - 841A - 2
US - 08 - 202 - 203 - 4
US - 08 - 202 - 203 - 203 - 19
US - 09 - 193 - 228 - 19
US - 09 - 193 - 228 - 19
US - 09 - 193 - 228 - 19
US - 09 - 348 - 817A - 5
US - 09 - 348 - 817A - 5
US - 09 - 348 - 817A - 5
US - 09 - 348 - 817A - 5
US - 09 - 348 - 843 - 4
US - 09 - 187 - 049 - 13
US - 09 - 187 - 003 - 4
US - 08 - 230 - 491A - 2
US - 08 - 230 - 491A - 2
US - 08 - 230 - 491A - 2
US - 08 - 230 - 491A - 2
US - 08 - 230 - 491A - 2
US - 08 - 230 - 491A - 2
US - 08 - 230 - 491A - 2
US - 08 - 230 - 391A - 236 - 6
                                                                                                                                                                                                                                                                                                                                                                                                                                     hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-937-236-5
US-08-569-214-2
US-09-058-489-36
                                                                                                                                                                                                                                                                                                                                                                                              212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Listing first 45 summaries
                                                                                                                                                                                                                                                                                   1 MVTPVTWMDNPIEVYVNDSV....

    protein search, using sw model

                                                                                                                                                 April 9, 2002, 16:57:24
                                                                                                                                                                                                                                                                                                                                                   Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                          US-09-490-700-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                       Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            108.5
93
93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83.5
82.5
82.5
82
82
82
82
82
82
                                                                                                                                                                                                                                                                 Perfect score:
                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        score gand is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred.
                                                                                                           OM protein
                                                                                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database
                                                                                                                                                      Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š.
```

0

32;

~

linear

TOPOLOGY:

```
966 HGFEVCNADAPVKWTLKLRFSDDQLRSILAKVCERAAVPRNWIQRLKKTLALGPTPPLKT 1025
                                                                                                                                                                                                                                                                                 SSIDQKLASQINDL--RQTVIWMGDRLDLEHHFQLQCDWNTSDFCIT-PQIYNESEHH-- 459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | : | : | : | | 846 NINEAVNFALPDWIFDDLESVRRYQRFRKPAVFSHDQLLITVSQOSQTIETAVWLEAAMQ 905
0.4 A---
                                                 241 WSGNQTLETRYRKPFYTIDLNSIL----TVPLQSCLKPPYMLVVGNIVIKPASQTITCEN 296
                                                                                                  721
                                                                                                                                                   CRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFI 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---PGTEAIAGVA 506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    906 EMVDREIAKRNALREIIPDLKEEVYDEDVAESHYICSHCTLFSYLGQLTSPKTDGVACLD 965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DGLANLN---PVTWIKTIR-STMIINLILIVVC------LFCLLLVCRCTPTAPKKT 553
                                                                                                                                                                                                                                                357 FTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLW-------NSQ 404
                                                                                               671 HSQKEEVEVEYGADVHSTTHGSALPTQETHPLSLYSRDKWNL--NNLPILPGS-----
                                                                                                                                                                                                                                                                                                                                                                                              786 RKAAPDLFETLPDLLFHLTTMMSPEKLKKEGVRVVACDQRANEFVVTFPKAYHSGFNHGL
                                                                                                                                                                                                                                                                                                                                                                                                                                               ------DMVRRHLQGRE-----DNLTLDISK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 2
US-07-862-021B-14
Sequence 14, Application US/07862021B
Sequence 14, Application US/07862021B
Sequence 14, Application US/07862021B
Sequence 15, Application US/07862021B
Sequence 16, Application US/07862021B
Sequence 17, Trive OF INVENTION: CLONING, EXPRESSION AND USES OF A TITLE OF INVENTION: CLONING, SECRETED PROTEIN, F-SPONDIN NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: 30 Rockefeller Plaza
STREET: 30 Rockefeller Plaza
  ----HLWRPDSEGLATSDYDPPTNGLARRVHGTDLVVSEDDVEREFWRL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,021B
FILING DATE: 19920405
CLASSIFCATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                               --LLQYIKSDIS-----PW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -- LKEQIFEASKAHLNLV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               568 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: New York
STATE: New York
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LRS 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        554 VTS 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1026
                                                                                                                                                                                                                                                                                                                                                  405
                                                                                                                                              297
                                                                                                                                                                                                                                                                                                 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   481
                                                                                                                                                                                                                                                                                                                                               δý
                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
```

```
22;
                                                                                                                                                                                                                                                                                           62 MPAVQNWLV-EVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKT---- 115
                                                                                                                                                                                                                                                                                                                                                                                                                        310 NEECE---PSSCIVTEWAEWEECSATCRMGMKKRH--RMIKMTPADGSMCKADTTEVEKC 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --PSAQVSPAVDSDLTESLDKHKHKKLQSFYLW-EWEE-----KGISTPRPKIISP--- 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     365 MMPECHTIPCVLSP------WSEWSDCSVTCGKGTRTRQRMLKSPSEL 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----VSGPEHPELWRLTVASHHIRIWSGNQTLETRYRKPFYTIDLNSILT 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 266 VPLQSCLKPPYMLVVGNIVIKP-----ASQTITCENCRLFTCIDSTFNWQHRILLVRAR 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 440 -ECNKCSGKGHMIRTRMITMEPQFGGAVCPETVQRKKCRLRKCQKSSGNERRHLKDAREK 498
                                                                                                                                                                                                                                                                                                                                        255 DDEASTCMMSEWITWSPCS----ASCGMGIEVRERYVKQFPEDGSLCKVPTEETEKCIV 309
                                                                                                                                                                                                                                                                                                                                                                                       ---CPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPR-GQFYHNCSGQTQSC 171
                                                                                                                                                                                                                                            203 IYSNWSPW----SACSSATCDKGKRMRQRMLKAQLDLSVPCPDTQDFEPCMG--PGCS 254
                                                                                                                                                  Gaps
                                                                                                                                                                                             14 VYVNDSVWVPGPTDDRCPAKPEEEG-----MMINISIGYHYP-----PICLGRAPGCL 61
                                                                                              3.1%; Score 93; DB 1; Length 568;
0.08%; Pred. No. 0.54;
.ve 38; Mismatches 148; Indels 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            499 RRSEKIKEDSDGEOYPVCKMKPWTA-----WTECTKFCGGGIQERFMTVKKRF 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------GVLNR----SKRF 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 107 GDCNEELELKQVEKCMLPECPISCELTEWSY----WS-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCT-US93-03164-14

Sequence 14, Application PC/TUS9303164

GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M
APPLICANT: Klar, Avihu
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E: Floppy disk
IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 401
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
                                                                                                                         20.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION:
                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19930402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA: APPLICATION NUMBER: FILING DATE:
; MOLECULE TYPE: protein US-07-862-021B-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM:
                                                                                                 Query Match
Best Local Similarity
Matches 86; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 1 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                       116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                               ò
                                                                                                                                                                                                                                            염
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δŽ
                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ω
```

```
NOVEL TYROSINE KINASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/08/323,474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TILLE OF INVENTION: Steven F.
TILLE OF INVENTION: NOVEL TYROSI
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corrections
STREET: 51 novements
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E: Immunex Corporation
51 University Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 8, Application US/08323474
Patent No. 5447860
GENERAL INFORMATION:
APPLICANT: Ziegler, Steven F.
                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 20.18
Matches 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: 373
CHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
US-08-202-841A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: Seattle
STATE: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 98101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-323-474-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                       . 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 MPAVQNWLV-EVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKT---- 115
                                                                                                                                                                                                                                                                                                                                                                                                                                     255 DDEASTCMMSEWITWSPCS-----ASCGMGIEVRERYVKOFPEDGSLCKVPTEETEKCIV 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --PSAQVSPAVDSDLTESLDKHKHKKLQSFYLW-EWEE-----KGISTPRPKIISP--- 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     365 MMPECHTIPCVLSP-------WSEWSDCSVTCGKGTRTRQRMLKSPSEL 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        266 VPLQSCLKPPYMLVVGNIVIKP-----ASQTITCENCRLFTCIDSTFNWQHRILLVRAR 319
                                                                                                                                                                                                                       Query Match 3.1%; Score 93; DB 5; Length 568;
Best Local Similarity 20.8%; Pred. No. 0.54;
Matches 86; Conservative 38; Mismatches 148; Indels 142; Gaps
                                                                                                                                                                                                                                                                                                                                                         203 IYSNWSPW-----SACSSATCDKGKRMRQRMLKAQLDLSVPCPDTQDFEPCMG--PGCS 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       116 ---CPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPR-GQFYHNCSGQTQSC 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   310 NEECE---PSSCIVTEWAEWEECSATCRMGMKKRH--RMIKMTPADGSMCKADTTEVEKC 364
                                                                                                                                                                                                                                                                                                              14 VYVNDSVWVPGPTDDRCPAKPEEEG-----MMINISIGYHYP-----PICLGRAPGCL 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------VSGPEHPELWRLTVASHHIRIWSGNQTLETRYRKPFYTIDLNSILT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        407 GDCNEELELKQVEKCMLPECPISCELTEWSY----WS-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               320 ------EGMWIPVSTDRPWEASPSIHILTEILK----GVLNR----SKRF 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      499 RRSEKIKEDSDGEQYPVCKMKPWTA-----WTECTKFCGGGIQERFMTVKKRF 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Hughes, Stephen H.
APPLICANT: Butrave, Pramod
APPLICANT: Sutrave, Pramod
APPLICANT: Pursel, Vernon
TITLE OF INVENTION: Enhancement of Musculature in Animals
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,841A FILING DATE: 23-FEB-1994 CLASSIFICATION: 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: Townsend and Townsend and Crew LLP STREEF: Two Embarcadero Center, Eighth Floor CITT: San Francisco STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 800
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,415
TITLE DATE: 03-DEC-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08202841A Patent No. 6218596 GENERAL INFORMATION:
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 568 amino acids
TYPE: AMINO ACID
                                                                                                                         MOLECULE TYPE: protein PCT-US93-03164-14
                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94111-3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-202-841A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                       음
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
```

```
/product= "OTHER"
/note= "Xaa = Trp in c-ski;
Xaa = Arg in v-ski"
FILING DATE: 02-UL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/373,864
FILING DATE: 30-UN-1989
ATTORNEY/AGENT INFORMATION:
NAME: Alicea, Hector A.
REGISTRATION NUMBER: 015280-170300US
TELEPONE: (415) 576-020
TELEPONE: (415) 576-020
TELEPONE: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 amino acids
TYPE: Amino acids
```

10; 319 FKTMLWDP--AGGSAVLQRQPDGNEVPSDP-----PASKKTKIDDSASQSPASTEKEKQ 370 -----VSGMSLRPRVNYLQD----FSYQRSLKFR-PKGKTCPKEIPKGSKN 126 259 SANWRSYILLSQDYTGKEEKARLGQLLDEMKEKFDYNNKYKRKAPRNRESPRVQLRRTKM 318 127 TEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAVDS---- 182 Gaps 2.9%; Score 87.5; DB 4; Length 750; 20.1%; Pred. No. 3.4; ive 32; Mismatches 82; Indels 6 183 ----DLTESLDK-----HKHKKLQSFYLW----EWEEKGISTPRPKII 217 47 YHYPPICLGRAPGCLMPAVQN------WLVEVPTVSPNSRFTYHM-

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```
TYPE: amino acid
STRANDEDNESS: un)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-469-537A-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-220-240A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            130 LVWEECVANSVVILQNNEFGTIIDXAPR------GQFYHNCSG----QTQSCPS 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         491 MDW-----STIVVDPSENVTLMNLRPKTGYSVRVQLSRPGEGGEGAMGPPTLMTTDCPE 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       234 ASHHIRIWSGNQTLETR--YRKPFYTIDLNSILT-----VPLQSC-----LKPPYML 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           279 VVGNIVIKPA-------SQTITCENCRLFTCIDSTFNWQH------RILLVRAR 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              635 -- AHVLLPPSGPPAPRHLHAQALSDSEIQL------TWKHPEALPGPISKYVVEVQVA 684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  72 VPTVSPNSRFTYHMVSGMSLRPRVNYLQD--FSYQRSLKFRPKGKTCPKEIPKGSKNTEV 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 AQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTV 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 579 -GFLLRLWDGTRGQERRENVSSPQARTALLTGLTPGTHYQLDVQLYHCTLLGPASPP--- 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         545 PLIQPWLEGWHVEGTDRLR-------VSWSLPLVPGPLVGD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                          2.9%; Score 87.5; DB 1; Length 1138; 19.9%; Pred. No. 7; tive 48; Mismatches 104; Indels 129;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   320 EG----MWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGL 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
APPLICANT: Maisonpierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
APPLICATION NUMBER: US 07/905,600
FILING DATE: 26-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/COCKET NUMBER: 2609
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: FSSLSED Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: 06-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 6
US-08-469-537A-98
; Sequence 98, Application US/08469537A
; Patent No. 5843749
                                                                                                                                                                                TELERAX: (206) 233-0644
TELEX: 756822
INPORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1138 amino acids
TYPE: amino acid
STRANDEDNESS: single
TYPE: MOLECULE TYPE: protein
US-08-323-474-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 2.9%
Best Local Similarity 19.9%
Matches 70; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STATE: NY COUNTRY: U. ZIP: 10591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2.9%; Score 87.5; DB 2; Length 1138;
19.9%; Pred. No. 7;
Live 48; Mismatches 104; Indels 129; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234 ASHHIRIWSGNQTLETR--YRKPFYTIDLNSILT------VPLQSC-----LKPPYML 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  279 VVGNIVIKPA-----RILLVRAR 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                72 VPTVSPNSRFTYHMVSGMSLRPRVNYLQD--FSYQRSLKFRPKGKTCPKEIPKGSKNTEV 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         130 LVWEECVANSVVILQNNEFGTIIDXAPR------GOFYHNCSG----QTQSCPS 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             491 MDW-----STIVVDPSENVTLMNLRPKTGYSVRVQLSRPGEGGEGAWGPPTLMTTDCPE 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 AQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTV 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            545 PLLQPWLEGWHVEGTDRLR------578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       579 -GFLLRLWDGTRGQERRENVSSPQARTALLTGLTPGTHYQLDVQLYHCTLLGPASPP--- 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       320 EG----MWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGL 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun STREET: 233 South Wacker Drive/6300 Sears Tower CITY: Chicago STATE: Illinois
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Matikainen, Marja-Terttu
APPLICANT: Partenen, Juha
APPLICANT: Makela, Tomi
APPLICANT: Korhonen, Jaana
APPLICANT: Korhonen, Jaana
TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR
TITLE OF INVENTION: TYROSINE KINASE AND USES THEREOF
NUMBER OF SEQUENCES: 5
                                                                           APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-0CT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
                                  USSN 08/406,247
                                                                                                                                                                                                                                   NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Sequence 5, Application US/08220240A
; Patent No. 5955291
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                    TELECOMMUNICATION INFORMATION
TELEPHONE: 914-345-7400
TELEFRAX: 914-345-7721
                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 98
SEQUENCE CHARACTERISTICS:
LENGTH: 1138 amino acids
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 19.9%
Matches 70; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Alitalo, Kari
APPLICANT: Matikainen, M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY: unknown
MOLECULE TYPE: protein
```

S

```
188 LDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWSGNQ-- 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    694 GRVSPSSSSDIAS-SLGESIESGPLSDILSDEESSMPLAGMKKYADEKSERASSSEKNES 752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      405 ----SIDQKLASQI----- 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      431 PP---DRSKLCLV--LOSSYPNSPSAASOSYECLHKVGNGNLENTVKFHIKEISSSLGRL 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---YQRSLKFRPKGKTCPKEIP------KGSKNTEVLVWEECVANSVVIL---- 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             246 -TLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCID 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             305 STFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIM 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          541 IEGPQTNSAST-SSLEPCNORSWNAKLQLQSETSS--SPAFTQSSESSVGSDNIMSPVPL 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLIAVTATAAVAGVALHSSVQS------VIEVNYWQKNSTRLWNSQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   598 LSKHKSKKGQA-----SP84 LSP84 LSP84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----QNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPA-----VDSDLTES----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 2.9%; Score 87; DB 2; Length 1075; Best Local Similarity 18.6%; Pred. No. 7.2; Matches 102; Conservative 70; Mismatches 171; Indels 206;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       50 PPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNY-LQDFS---
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Munice
APPLICANT: Spaulding, Vikki
APPLICANT: Spaulding, Vikki
APPLICANT: Spaulding, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,228
                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: Genetics Institute, Inc. STREET: 87 CambridgePark Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-824
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1075 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: amino acid STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                 Cambridge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      us-08-993-228-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              234 ASHHIRIWSGNQTLETR--YRKPFYTIDLNSILT-----VPLQSC-----LKPPYML 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         279 VVGNIVIKPA------SQTITCENCRLFTCIDSTFNWQH-------RILLVRAR 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               72 VPTVSPNSRFTYHMVSGMSLRPRVNYLQD--FSYQRSLKFRPKGKTCPKEIPKGSKNTEV 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       130 LVWEECVANSVVILONNEFGTIIDXAPR------GQFYHNCSG----QTQSCPS 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 AQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTV 233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              579 -GFLLRLWDGTRGOERRENVSSPQARTALLTGLTPGTHYQLDVOLYHCTLLGPASPP--- 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           635 -- AHVLLPPSGPPAPRHLHAQALSDSEIQL-----TWKHPEALPGPISKYVVEVQVA 684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          545 PLLOPWLEGWHVEGTDRLR------VSWSLPLVPGPLVGD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         320 EG----MWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGL 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 2.9%; Score 87.5; DB 2; 18est Local Similarity 19.9%; Pred. No. 7; Matches 70; Conservative 48; Mismatches 104;
                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/F193/00006
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/817,800
FILING DATE: 09-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/167,453
FILING DATE: 15-DEC-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29151/31958
                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/220,240A
FILING DATE: 29-MAR-1994
CLASSIFICATION: 435
       Unites States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 19, Application US/08993228
Patent No. 5976838
GENERAL INFORMATION:
APPLICANT: Jacobs, John M.
APPLICANT: LaVallie, Edward R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Gass, David A.
REGISTATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 2915;
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
INFORMATION FOR SEO ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 1138 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-993-228-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-220-240A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

28;

404

```
226 KEH-----SFEVDVWSIGCIMYTLLVGKPPFETSCLKETYLRIKKNEYSIPKHINP 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                              --- DXAPRGQFYHNCSGQTQSCP 172
                                                                                                                108 LAHQHVVGFHGFFEDNDFVFVVLELCRRRSLLEPHKRRKALTEPEARYY - - LRQIVLGCQ 165
                                                                                                                                                                                                     166 YLHRNRVIHRDLKLGNLFLNEDLEVKIGDFGLATKVEYDGERKKTLCGTPNYIAPEVLSK 225
                                                                                                                                                                                                                                                    223 PEHPELWRLTVASHHIRIWS-----GNQTLET-----RYRKPFYTI--DLNS 262
                                                                                                                                                                                                                                                                                                                                                                                     277 VAASLIQKMLQTDPTARPTINELLNDEFFTSGYIPARLPITCLTIPPRFSIAPSSLDPSN 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TGEVVDCHLSDMLQQLHSVNASKPSERGLVRQEEAEDPACI-----PIFWVSKWVDY 417
                                                                                                                                                             SAQVSPAVDSDLTES---LDKHKHKKLQSFYL-----WEWEEKGISTPRPKIISP--VSG 222
                                                                                                                                                                                                                                                                                                                                             263 ILTVPLQSCLK--PPYMLVVGNIV-----IKPASQTITCENC-RLFTCIDSTFNWQH 311
                                                                                                                                                                                                                                                                                                                                                                                                                                    312 RILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTA 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           372 TAAVAGVALHSSVQSVNFVNYWQKNSTRLMNSQSSIDQKLASQINDLRQTVIWMGDRLDL 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     432 EHHFQL---QCD-----WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLD---- 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             418 SDKYGLGYQLCDNSVGVLFNDSTRLI---LYNDGDSLQYIER---DGTESYLTVSSHPNS 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----ISKLK------EQIFEASKAHLNLVPGTEAIAGVADGLANLNPV-TWIKTIRSTM 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: COLEMAN, TIM
APPLICANT: LAWRENCE, DANIEL
TITLE OF INVENTION: BRAIN-ASSOCIATED INHIBITOR OF
TITLE OF INVENTION: TISSUE-TYPE PLASMINGEN ACTIVATOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E: HUMAN GENOME SCIENCES, INC 9410 KEY WEST AVENUE
                                                                    135 CVANSVV----ILQNNEFGTII-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/08/948,997 FILING DATE: Oct-10-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF336
TELECOMMUNICATION:
TELECHONE: (301) 309-8504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 5, Application US/08948997; Patent No. 6008020
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: A. ANDERS BROOKES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: HASTINGS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ROCKVILLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   526 IINL 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             522 ILHL 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-948-997-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET:
                                                                                                                                                           173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        366
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      염
                                                                 ŏ
                                                                                                            q
                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32;
                                                                                         :| | | : | | : | | : | | 313 DDLKLYLETHLSFKL--NVDSHCALKEAVEEEGHQLLELIASHKAGLKDMLRMIASQWKE 870
---RLDLEHHFQLQCDWNTSDFCITPQIYNESEHH-WDMVRRHLQGREDNLTLDISKLKE 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGMSLRPRVNYLQDFSYQRSLKFRP-----KGKTCPKEI-----PKGSKNTEVLVWEE 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12 RAPADPGKAGV-----PGVAAPGAPAAAPPA----KEIPEVLVDPRSRRRY--- 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29 RCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNWLVEVP--TVSPNSRFTYHMV 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        221; Indels 193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 4; Length 603;
                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Strebhardt, Klaus; Rubsamen-Waigmann, Helga APPLICANT: Holtrich, Uwe TITLE OF INVENTION: CLONING OF A MEMBER OF THE SERINE-TITLE OF INVENTION: THREONINE-KINASE FAMILY NUMBER OF SEQUENCES: 7 CORRESPONDENCE ADDRESS: ADDRESS: SPRUNG HORN KRAMER & WOODS STREET: 660 White Plains Road
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
MEDIUM TYPE: storage
COMPUTER: NEC Powermate SX-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2.8%; Score 85.5; Di
19.2%; Pred. No. 3.9;
:ive 74; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 9516-KGB
TELECOMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/601,014
FILING DATE: 23-FEB-1996
APPLICATION NUMBER: PCT/EP94/02863
FILING DATE: 30-AUG-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 30-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4329177
                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/09198122 Patent No. 6180380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FILING DATE: 30-AUG-1993 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : 603 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 19.2
Matches 116; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: Tarrytown STATE: New York
                                                                                                                                                        484 QIFEASKAH 492
                                                                                                                                                                                                       871 LQRQIKRQH 879
                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A. ZIP: 10591-5144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                         RESULT 9
US-09-198-122-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-198-122-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                 õ
                                                                                                              g
                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
```

```
11 DGKSYQVPMLAQLSVFRSGSTKTPNGLWY-NFIELPYHGESISMLIALPTESSTPLSAII 269
                                                                                                     | | | : : | | | | : : | | | 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            : :: | | : : | | : 438 FKVNVKVPPVPLTAPRLLAKQSRQLVVSPLVSFGGDGPISSVRLHYRPQDSMITW---- 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  138 NSVVILQNNEFGTIIDXAPR------GQFYHNCSG----QTQSCPSAQVSPAVD 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PQIYNESEHHW--DMVRRHLQGREDNLT-LDISKLKE-----QIFEASKAHLNLVPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.8%; Score 85; DB 2; Length 1135;
18.1%; Pred. No. 13;
ive 59; Mismatches 107; Indels 156;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E: Regeneron Pharmaceuticals, Inc. 777 Old Saw Mill River Road
                                                                                                                                                                                                                                                                                                                                                                                                                TYROSINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-OCT-1993
ATTONNEY, AGGNT INFORMATION:
NAME: Kempler, Ph. D., Gall M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION 1NFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/469,537A FILING DATE: 06-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              103 YQRSLKFRPKGKTCPKEIPKGSKNTEV---
                                                                                                                                                                                                                                                                                                               Sequence 97, Application US/08469537A Patent No. 5843749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: Malsonpierre, et al
TITLE OF INVENTION: EHK AND RO
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: REGENETON Pharmac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRA.

ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
TOMPOTER: IBM COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
LENGTH: 1135 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 2.8%
Best Local Similarity 18.1%
Matches 71; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; MOLECULE TYPE: protein US-08-469-537A-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET: 777 ord
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                             499 TEAI 502
                                                                                                                                                                                                         330 SESL 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: NY COUNTRY: U
                                                                                                                                                                                                                                                                          RESULT 12
US-08-469-537A-97
                                                                          450
                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pp
                                                                       δ
                                                                                                                                                             ŏ
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Hastings et al.
APPLICANT: Hastings et al.
TITLE OF INVENTION: Brain-Associated Inhibitor of Tissue-Type Plasminogen
TITLE OF INVENTION: Activator
FILE REFERENCE: PF336D1
CURRENT APPLICATION NUMBER: US/09/348,817A
CURRENT FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 08/948,997
PRIOR FILING DATE: 1997-10-10
PRIOR FILING DATE: 1996-10-11
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                :| :| | :::: : | |: : | |
91 VLKKINKAIVSKRNKDIVIVANAVFVRNGFKVEVPFAARNKEVFQCEVQSVNFQDPASAC 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                        151 DAINFWVKNETRGMIDNLLSPNLIDSALTKLVLVNAVYFKGLWKSRFQPENTKKRTFVAG 210
                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                 ------ TWNSQ------- 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --SSIDQKLASQINDLRQ-----TVIWMGDRLDLEHH-----FQLQCDWNTSDFCIT 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                211 DGKSYQVPMLAQLSVFRSGSTKTPNGLWY-NFIELPYHGESISMLIALPTESSTPLSAII 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     270 PHISTKTINSWANTMVPKRAQLVLPKFTALAQTDLKEPLKALGITEMFEPSKANFAKITR 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              340 ILTEILKGVLNRSKRFIFTL---IAVIMGLIAVTATAAVAGVALHSSVQSVNF----- 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      390 --VNYWQKNSTR----- 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           405 --SSIDQKLASQINDLRQ-----TVIWMGDRLDLEHH-----FQLQCDWNTSDFCIT 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PQIYNESEHHW--DMVRRHLQGREDNLT-LDISKLKE-----QIFEASKAHLNLVPG 498
                                                                                                                                                                                                                                                                                               82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82;
                                                                                                                                                                                                                                                                                                                                      340 ILTEILKGVLNRSKRFIFTL---IAVIMGLIAVTATAAVAGVALHSSVQSVNF
                                                                                                                                                                                                                                                     Length 397;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 4; Length 397;
                                                                                                                                                                                                                                                                                            76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                   DB 3;
                                                                                                                                                                                                                                                2.8%; Score 85; DB 3
20.5%; Pred. No. 2.1;
tive 36; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 2.8%; Score 85; DB (Best Local Similarity 20.5%; Pred. No. 2.1; Matches 50; Conservative 36; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                            --VNYWQKNSTR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 5, Application US/09348817A Patent No. 6191260
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; ORGANISM: Rattus norvegicus US-09-348-817A-5
                                                                                                                                                                                                                                   Query Match
Best Local Similarity 20.5%
                                                                                        TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-997-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :1::
330 SESL 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    499 TEAI 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
US-09-348-817A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
```

g

ð

ð g

g

ò

ò

a

ò

ò g ð

õ a

```
19;
```

us-09-490-700-38.rai

	λŏ	302 CI
OY 182 SDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIW 241	. 5	376
Db 552 SWNVEGPDRLRVSWSLPSVPLSGDGFLLRLW 582		0/0
Qy 242 SGNQTLETRYRKPFYTIDLNSILTVPLQSCL 272	Qy	340
:	qq	427 OT
Ov 273 KPDVMIVVGNIVIKPASOFITCENCRIFFCIDSTFWWOHRILLVRAREG- 321	Qy	378 VA
643 APRHIRACALSOSETERMWOHDEAPDEDTSKYTVETOVAGG	qa	486 AA
TO C.C.	09	433 HH
5.2	q _Q	541
GOS GUFÇMMDVDAFEELSTIVKG-LNAS	Qy	493 LN
Qy 379 ALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKL 411 : : : : : : :	q a	585
RESULT 13	RE CIS	RESULT 14 US-09-376-78
US-09-111-085-2	••	Sequence
; sequence 2, Application US/USILIUSS ; Patent No. 6100034		Patent NO. GENERAL IN
; GENERAL INFORMATION:		APPLICANT
; APPLICANT: Welss, Robin A		APPLICAN
; TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope		TITLE OF
; FILE REFERENCE: 4238/75168		TITLE OF
; CURRENT APPLICATION NUMBER: US/09/111,085	•••	FILE REFE
CURRENT FILING DATE: 1998-07-07 ; EARLIER APPLICATION NUMBER: GB 9710154.7		CURRENT
; EARLIER FILING DATE: 1997-05-16	•••	EARLIER A
SOFTWARE: PatentIn Ver. 2.0		NUMBER OF
; SEQ ID NO 2 ; LENGTH: 660		SOFTWARE: SEO ID NO
; TYPE: PRT	•••	LENGTH:
) ONCHRISH: PULCTURE TELLOVILUS US-09-111-085-2		ORGANISM
	•••	FEATURE:
2.88		OTHER IN
9; Conservative 74; Mismatc	so l	US-09-376-78
Qy 6 TWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPI 52		
Db 82 TWWPELYVCLRSVIPGLNDQATPPDVLRAYGFYVCPGPPNNEFYCGNPQD 131		Query Matc Best Local
Qy 53 CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLK 108		Matcnes ,
DD 132 FFCKQWSCITSNDGNWKWP-VSQQDRVSYSFVNNPTSYNQFNYGHGRWKDWQQRVQKD 188	δλ	6 TK
	qa —	82 TW
100 total control of the control of	Qy	53 CI
TOO MUNICIPALITY TO THE TENGENT TO T	qa 	132 FF
QY TOI IMNCSG	QY	109 FF
200 IIOOOOMAANOO EITKEIKEELEMINE KATTO MAGEELEMOOF EIGEN. 308 CICMDDDEXIICONSC DEEDDE MAN CHIIDITECKAAAN EMBVEREDENIID I	qa	189 VR
	Qy	161 YH
THE PROPERTY OF THE PROPERTY O	qa	230 YY
ZOI NSILIVELQSCINTERNACIONI NI :	ον	208 GI
DD 316 NSTTPEATSSCWLCLASGPPYYEGMARGGKFNVTKEHRDQCTWGSQNKLTLT-EVSGKGT 374		t

```
36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               f: Andersson, Goran K.
INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
6261806
                                                                                                                                                                                                                                           276 --PSPNPSDYNTTSGSVPTEP------NITIKTG-AKLFSLIQGAFQAL 315
----- DSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIH----- 339
                                     GMVPPSHQHLCNHTEAFN------RTSESQYLVPGYDRWWACNTGLTPCVSTLVFN 426
                                                                                                                                 ALHSSVQSV-----NFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRLDLE 432
                                                                                                                                                                                                                                                                                                            FQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAH 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNY----LQDFSYQRSLK 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CKQWSCITSNDGNW--KWP-VSQQDRVSYSFVNNPTSYNQFNYGHGRWKDWQQRVQKD 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RPKGKTCPK-----EIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQF 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INCSG------QTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEK 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STPRPKIISPVSG--PEHPELWRLTVASHHIRIWSGNQTLETRYRKPFYTID-----L 260
                                                                                                    ----ILTEIL-----KGVL-----NRSKRFIFTL-IAVIMGLIAVTATAAVAG 377
                                                                                                                                                                                                                                                                                                                                               MDNPIEVYVNDSVWVPGPTDDR------CPAKPEEEGMMINISIGYHYPPI 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WFORMATION: Description of Artificial Sequence:PERV-A WFORMATION: polypeptide sequence taken from GenBank Accession WFORMATION: No. 6261806 Y12238 for comparison.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RNKQISCHSLDLDYLKİSFTEKGKQENIQKWVNGISWGIV------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2.8%; Score 84; DB 4; Length 660;
19.8%; Pred. No. 6.6;
tive 74; Mismatches 226; Indels 222;
                                                                                                                                                                                                                                                                                                                                                                                                                                                ILVPGTEAIAGVADGLANLNPVTWIKTIRSTM---IINLILIVVCLFCLL 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PETINGENTION: USE
PERENCE: 61750-267

RAPLICATION NUMBER: US/09/376,781

FILLING DAPE: 1999-08-18

RAPLICATION NUMBER: 60/097,015

RILING DAPE: 1998-08-18

RELING DAPE: 1998-08-18

RELING DAPE: 1998-08-18

OF SEO ID NOS: 33

RE: PatentIn Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Application US/09376781 6261806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Banerjee, Papia T.
Patience, Clive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
```

21;

Gaps

53; Mismatches 132; Indels 138;

82; Conservative

Matches

```
CI----- DSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIH----- 339
                                                                                                                        375 CIGMVPPSHQHLCNHTEAFN-----RTSESQYLVPGYDRWWACNTGLTPCVSTLVFN 426
                                                                                                                                                                .....ILTEIL-----KGVL-----NRSKRFIFTL-IAVIMGLIAVTATAAVAG 377
                                                                                                                                                                                                                                                   378 VALHSSVQSV-----NFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRLDLE 432
                                                                                                                                                                                                                                                                                          486 AALITGPQQLEKGLSNLHRIVTEDLQALEKSVSNLEESLTS----LSEVVLQNRRGLDL- 540
                                                                                                                                                                                                                                                                                                                                      433 HHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAH 492
                                                                                                                                                                                                                                                                                                                                                                     -----LFLKEGGLCVA--LKEECCFYVD----HSGAIRDSM----SKLRERLERRRRER 584
  ----TITCENCRLFT 301
                                     316 NSTIPEATSSCWLCLASGPPYYEGMARGGKFNVTKEHRDQCTWGSQNKLILL - EVSGKGT 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----EADQGWFEGWFNRSP--WMTTLLSALTGPLVVLLLLLTVGPCLI 626
                                                                                                                                                                                                                                                                                                                                                                                                                     493 LNLVPGTEAIAGVADGLANLNPVTWIKTIRSTM---IINLILIVVCLFCLL 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 13, Application US/09187049
Patent No. 6117666
CENERAL INFORMATION:
CENERAL INFORMATION:
TITLE OF INVENTION: PLASTID PROTEOLYTIC PROCESSING ENZYME TITLE OF INVENTION: THAT CLEAVES PRECURSOR POLYPEPTIDES NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS HOFER GILSON & LIONE
STREET: P.O. Box 10395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/187,049
FILING DATE:
NSILTVPLQSC-----LKPPY---MLVVG--NIVIKPASQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7814/16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/695,177
FILING DATE:
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814,
TELECOMUNICATION INFORMATION:
TELEPHONE: 312 321-4200
TELEFAX: 312 321-4209
INFORMATION FOR SEQ ID NO: 13:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 1259 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STREET: P.O. F
CITY: Chicago
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -09-187-049-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-187-049-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                 302
                                                                                                                                                                    340
261
                                                                                                                                                                                                                                                                                                                                                                                                                                                               585
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
ò
                                       8
                                                                                 ð
                                                                                                                 g
                                                                                                                                                                ò
                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

Length 1259;

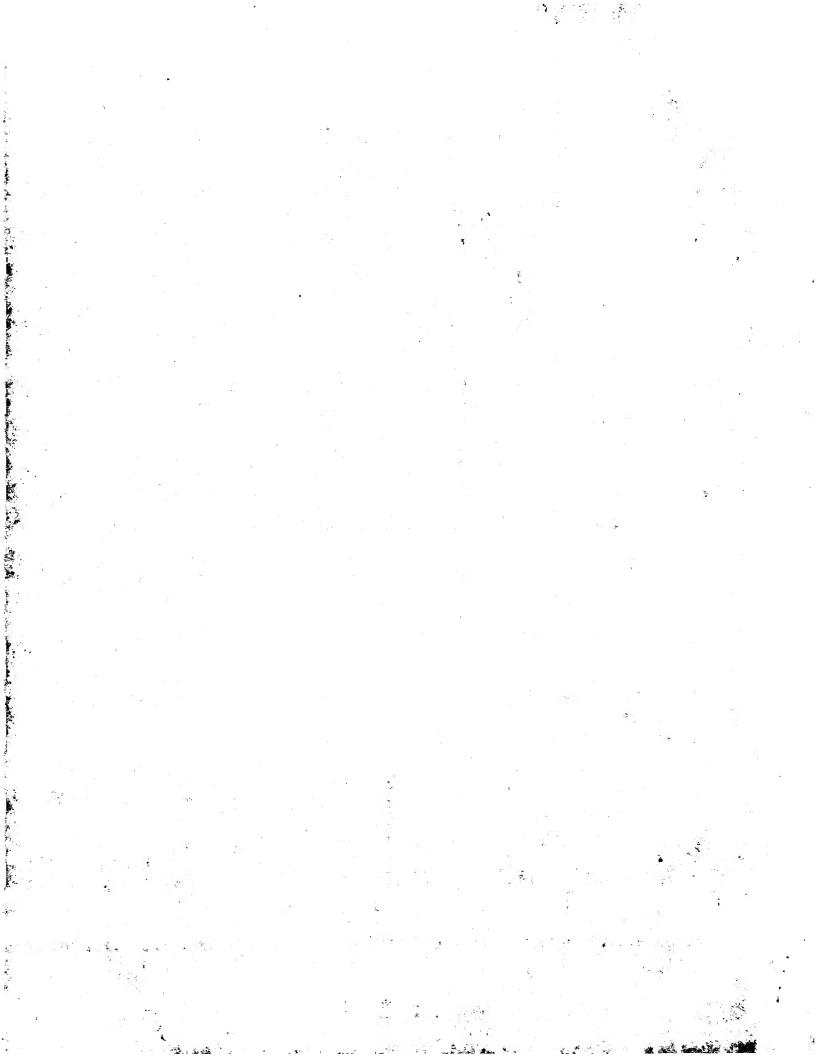
Score 84; DB 3; Pred. No. 20;

2.8%;

Query Match Best Local Similarity

```
LOS-----QEVFLNDTDERACAY-----IAGPA-PNRWGFTA------DGNDLL 1028
                                                                                                                                                                                                                                                                                                                                                                          882 RQVYL---SYYRSIPKSLERSTAHKLMVAMLDGDERFTEPTPSSLENLTLQSVKDAVMNQ 938
                                                                                                                                                                                                                                                         188 LDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWSGNQTL 247
                                                                                                                                                                                                                                                                                                                                                                                                                                295 ENCRLFTCIDSTFNWQHRI-----LLVRAREGMWIPVSTDRPWEASPSIHILTEILKGV- 348
                                                                                                                                                                    135 CVANSVVILQNNEF-----GTIIDXAPRGQFYHNCSGQTQSC-PSAQVSPAVDSDLTES 187
                                                                                                                                                                                                            939 FVGNNMEVSIVGDFTEEEIESCILDYLGTAQATGNFKNQQQIIPPTFRLSPS-----S 991
                                                                                                                                                                                                                                                                                                                                           248 ETRYRKPFYTIDLNSILT------VPLQSCLKPPYMLVVGNIVIKPASQTITC 294
                                                                                   RVNYLQDFSYQRSLKFRPKGKTCPK-------EIPKGSKNTEV----LVWEE 134
                                 CVNNQINCSLESTEEF----ISLEFRFTLRNNGMRAAFQLLHMVLEHSVWSDDALDRA 881
CLGRAPGCLMPAVQNWLVEVPTVSPNSRFT------YHMVSGMSL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |:|:| | |:|:|
1135 SNGITVRELDRAKR---TLLMRHEAEIKSNAYWLGLLAHLQSSSV 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----LNRSKRFIFTLI-----AVIMGLIAVTATAAV 375
                                                                                                                                                                                                                                                                                                  992
                                        828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      349
23
                                                                                 94
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                οg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
ò
                                    qq
                                                                                   ò
                                                                                                                       g
                                                                                                                                                                    ^{\circ}
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

Search completed: April 9, 2002, 16:57:28 Job time: 247 sec



```
April 9, 2002, 16:58:10; Search time 39.41 Seconds (without alignments) 1084.342 Million cell updates/sec
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                        OM protein - protein search, using sw model
                                                                                                                                                       Run on:
```

US-09-490-700-38 3011 1 WVTPVTWMDNPIEVYVNDSV......VCRCTPTAPKKTVTSRTGHE 561 Title: Perfect score: Sequence: Scoring table:

219241 seqs, 76174552 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

219241 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES æ

		Describtion	retrovirus-related	env polyprotein -	env polyprotein -	env polyprotein -	env polyprotein pr	env polyprotein pr	CO.	~	7	~	7	à	7	hilin-2 (protein -	CL2BB protein - ra	CL2AA protein - ra	CL2AC protein - ra		CL2BC protein - ra	alpha-latrotoxin r	ei	latrophilin-2 (spl	a	~	hilin-2 (_	CL1BA protein - ra	CL1AB protein - ra
	ţ	TT	VCHUER	VCMVMM	VCMVM	S26388	VCMVJA	VCMSIA	T18385	T18387	T18386	T18388	T18381	T18383	T18382	T18384	T17158	T46611	T17157	T17159	T17160	T17185	T14324	VCLJC6	T18375	T18379	T18377	T18380	T17138	_	117145
	6	9 ;	4	Н		~	-	-	~	~	N	N	(7	7	7	7	~	7	N	~	~	~	~	П	7	~	7	7	~	7	7
	4	Leng Ci	584	689	688	688	615	584	1420	1435	1463	1478	1407	1422	4	1465	1420	マ	4	1463	4	1478	1487	964	1354	1369	1397	1412	1466	1471	1510
p p	Query	Marcii	8.06	13.2		12.9		5.0	4.4	4.4	4.4	4.4	4.2	4.2	4.2	4.2	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	3.9	9. 6.	3.7	3.7	3.7
	9	arose	2735	398	393		340.5	152	131	131	131	131	125.5	125.5	125.5	125.5	118	118	-	118	118	118	7	116.5	116	116	116	_		111.5	111.5
	Result		1	7	m	4	z,	9	7	60	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

CLIBB protein - ra latrophilin-2, spl	<pre>latrophilin-2, spl latrophilin-2, spl</pre>	latrophilin-2 (spl env polyprotein pr	latrophilin-1, bra	Latrophilin-1, bra env polyprotein pr	env polyprotein pr	env polyprotein pr	env polyprotein -	hypothetical prote	toucan gene protei	env polyprotein -
T17156 T18301	T18367 T18366	T18370 VCLJE3	T18411	TI8413 VCLJEV	VCLJ22	VCLJEW	VCLJG3	S56268	T13806	VCFVER
0.0	00	77	~ (7	н,	٦.		7	~	7
1515	1356	1399 859	1467	14 / 2 859	823	829	881	787	2176	603
3.7	3.7	3.7	9.0	9.8 9.9	9.0	ص ۱۰ ۱۰	. w . s.	3.5	3.5	3.4
111.5 110.5	110.5	110.5 108	107.5	107.5	107	107	105	104	104	103.5
30	332	34	36	38	33	4 €	42	43	44	45

ALIGNMENTS

the mou

CHUER CCHUER CCHUER CCHUER CCHUER CCHUER CCESSIO CONTONIO	RESULT 1 VCHUER VCHUER VCHUER VCHUER C; Species: 28- C; Accession R; Accession A; Status: C A; Molecule A; Reference A; Residues: A; Gene: env C; Genetics: A; Gene: env C; Covertics: A; G	RESULT 1 VCHUER retrovirus-related env polyprotein pseudogene - human C;Species: Homo sapiens (man) C;Date: 28-Dec-1987 #sequence_revision 04-Jan-1996 #text_change 14-May-1999 C;Date: 28-598, 1986 R;Ono, M.; Yasunaga, T.; Miyata, T.; Ushikubo, H. J. Virol. 60, 589-598, 1986 A;Title: Nuclectide sequence of human endogenous retrovirus genome related to the A;Reference number: A93023; MUID:87036922 A;Accession: E24483 A;Accession: E24483 A;Accession: E24483 A;Accession: L2483 A;Accession: L2483 A;Accession: L2484 A;Accessio	ogene 584;	PUTWMDNPIEVYVNDSTWOLEY, THE THE THE TOTAL	AVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKG 123	LTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWSG 243	DSTFWNQHRILLVRAREGGWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVI DSTFWNQHRILLVRAREGGWIPVSTDRPWEASPSIHILTEVLKGVLNRSKRFIFTLIAVI BSTFWWQHRILLVRAREGGWIPVSMDRPWEASPSVHILTEVLKGVLNRSKRFIFTLIAVI MGLIAVTATAAVAGVALHSSVQSVNRVNVWQKNSTRLWNSQSSIDQKLASQINDLRQTVI
		CHUER CHUER etroviru Species) Date: 23) Accessic (Ono, M. VIII) 1, Title: 1), Teferen 3, Accessic Status: 3, Status: 3, Status: 3, Moleculus: 3, Moleculu	Cross-refer Genetics: Gene: env Gene: env Keywords: c	17 4 19 19 19 19 19 19 19 19 19 19 19 19 19	y 64 bb 61 y 124 bb 121		

ò

φ

```
A;Gene: env
C;Superfamily: type A retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein
F;1-98/Domain: leader peptide #status predicted <LPT>
F;9474/Aproduct: coat protein gp36 #status predicted <GPI>
F;475-688/Product: coat protein gp36 #status predicted <GPI>
F;127,143,297,498,557/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N;Contains: coat protein gp36; coat protein gp52
C;Species: mouse mammary tumor virus, MMTV
C;Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 04-Dec-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         G-----KGDKRRMWELWLHTLGNSGA---NTKLVPIKKKLP--PKYPHCQIAFKKDA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --SGQTQSCPS-----SDLT 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 FWEGDESAPPRWLPCAFPDKGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYT 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ESLDKHKHKKLQSFYL---WEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWS 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RLTNCLDSS-AYDYAAIIVKRPPYVLLPVDIGDEPWFDDSAIQTFRYATDLI----RAK 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLAS 413
                               QINDLRQTVIWMG-DRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQG--R 470
                                                                 EDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLI 530
                                                                                                                                                                                      PVGWGSTDPIRVLTNQTMYLGG-----SPDFHGFR-NMSGNVHFEGKSDTLPICFSF 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 -LGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            300 PPVNKEVHRWYEAGWVEPTWFWEN---SPKDPNDRDFTALVPHTELFRLVAASRHLIL-- 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------KRPGFQEHEM----IPTSACVIYPYAILLGLPQLIDIEKRGSTFHISCSSC 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   298 RLFTCIDSTFNWQHRILLVRAREGMWIPVST-DRPW---EASPSIHILTEILKGVLNRSK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PVTW-MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGGMMINISIGYHYP-----PIC--- 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNOTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVG-----NIVIKPASQTITCENC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNNLHRNVTLALSEQRIIDLKLEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 393; DB 1; Length 688;
Pred. No. 8.7e-24;
9; Mismatches 219; Indels 152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C;Accession: A03972
R;Redmond, S.M.S.; Dickson, C.
Rsmbo J. 2, 125-131, 1983
A;Title: Sequence and expression of the mouse mammary tumour virus A;Reference number: A03972
A;Accession: A03972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 env polyprotein - mouse mammary tumor virus (strain GR)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SFSTPTGCFQVDKQVFLSDTPTVDNN-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13.1%; 23.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 145; Conservative
                                                                                                                                                                                                                                                                       LIVVCLFCLLLVCRC 545
                                                                                                                                                                                                                                                                                                      |:|:| :|
LLVIVLMIFPIVFQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-688 <RED>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Genetics:
                                                                                      516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                456
                               414
                                                                                                                                                 471
                                                                                                                                                                                                         575
                                                                                                                                                                                                                                                                       531
                                                                                                                                                                                                                                                                                                                               634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                Q
                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                        env polyprotein - mouse mammary tumor virus
NiContains: coat protein gp52
C; Species: mouse mammary tumor virus, MMTV
C; Species: mouse mammary tumor virus, MMTV
C; Species: mouse mammary tumor virus, MMTV
C; Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C; Accession: D26795
R; Moore, R: Dixon, M: Smith, R: Peters, G.; Dickson, C.
J. Virol. 61, 480-490, 1887
A; Title: Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: the particle complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: the particle type: DNA
A; Accession: D26795
A; Molecule type: DNA
A; Accession: D26795
A; Molecule type: DNA
A; Mo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25;
                                                                                                               EQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCLLLV 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         G------KGDKRRMWELWLTTLGNSGA---NTKLVPIKKKLP--PKYPHCQIAFKKDA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :: | | | | | : : : : | | | | 300 PPVNKEVHRWYEAGWVEPTWFWEN---SPKDPNDRDFTALVPHTELFRLVAASRHLIL-- 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     354 RFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLAS 413
WMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLK 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --SGQTQSCPS-----SDLT 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RLTNCLDSS-AYDYAAIIVKRPPYVLLPVDIGDEPWFDDSAIQTFRYATDLI-----RAK 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PVGWGSTDPIRVLTNQTMYLGG-----SPDFHGFR-NMSGNVHFEGKSDTLPICLSF 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 PVTW-MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYP-----PICL-- 54
                                  GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ESLDKHKHKKLQSFYL---WEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 FWEGDESAPPRWLPCAFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKNLFGNYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVG-----NIVIKPASQTITCENC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RLFTCIDSTFNWQHRILLVRAREGMWIPVST-DRPW---EASPSIHILTEILKGVLNRSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 146; Conservative 100; Mismatches 217; Indels 152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13.2%; Score 398; DB 1; 23.7%; Pred. No. 3.4e-24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SESTPTGCFQVDKQVFLSDTPTVDNN------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                        543 CRCT 546
                                                                                                                                                                                                                                                                                                   CRCT 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    113
434
                                                                                                                  483
                                                                                                                                                                             480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                 g
                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

a á ð

δý

ò

25;

```
VCMVJA

VCMVJA

VCMVJA

VCMVJA

N'Alternate names: coat polyprotein

N'Alternate names: coat polyprotein

N'Alternate names: coat polyprotein

N'Alternate names: coat protein gp52

Cycotesins: coat protein gp52

Cycotesion: Britan

Cybate: 31-Mar-1993

Received: Species: sheep pulmonary adenomatosis virus

Cybate: 31-Mar-1993

Rydression: E42740

Rydression: E42740

Ayritle: Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenou

Ayritle: Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenou

Ayritle: Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenou

Ayritle: Nucleotide sequence of the jaagsiekte retrovirus, an exogenous

Ayrotecsion: E42740

Ayrotecin e42740

Ayrotecin type: genomic RNA

Ayrotecine type: genomic RNA

Ayrotecine type: genomic RNA

Ayrotecine type: genomic RNA

Ayrotecine type: genomic RNA

Cycoss-references: GB:M80216; NID:g331338; PIDN:AAA89184.1; PID:g331342

Cyconetics:

Ayrotecine type: genomic RNA

Cyconetics:

Ayrotecine type A retrovirus env polyprotein; transmembrane predicted (SDP)

Cycomath: signal sequence #status predicted (SDP)

F:379-402.Dommain: transmembrane #status predicted (CDP)

F:379-402.Dommain: intracellular #status predicted (TNT)

F:379-402.Dommain: intracellular #status predicted (TNT)

F:375-402.571/Dommain: transmembrane #status predicted (TNT)

F:555-571/Dommain: transmembrane #status predicted (TNT)

F:5108,127,178,219,275/Binding site: carbohydrate (ASD) (covalent) #status predicted

F:108,127,178,219,275/Binding site: carbohydrate (ASD)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1: : | : ||||| : ||:|
92 MIQSLGWDREIVPVYVNDTSLLGGKSD--IHISPQQ----ANISFYGLTTQY-PWCFSYQ 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGK 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVN 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                         SERVIQELTNLIGDMSKQHIDTVDLSGLAQSFANGVKALNPLDWTQYF-IFIGVGAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTP-----RPK1ISPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------FLWDWAPGG--TPDFPEYRGQHPPIFSVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGPEH-PELWRLTVASHHIRIWSGNQTLETRYRKP-----FYTIDLNSILTVPLQSCLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TAPIYQTELWKLLAAFGH-----GN----SLYLQPNISGTKYGDVGVTGFL-YP-RACVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       274 PPYMLVVGNIVIKPASQT--ITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINIS---IGYHYPPICL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 115 TCPKEIPKGSKNTEVLV----WEECVANSVVILQ-NNEFGTIIDXAPRGQFYHNCSGQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGPLDIPFCDKHLSIGIGIDTPWTLCRARVASVYNINNANAT ...........
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 11.3%; Score 340.5; DB 1; Best Local Similarity 23.8%; Pred. No. 1.4e-19; Matches 139; Conservative 100; Mismatches 202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               145 SQHPHCIQVSAD---ISYPRVT-----ISGID--
                                                                                                                                                                                                                            648
                                                                                                                                                                          531 LIVVCLFCLLLVCRC 545
                                                                                                                                                                                                            |:|: | :|: |
LLVIVLMIFPIVFQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      221
                          516
                                                                         471
                                                                                                                           575
                                                                                                                                                                                                                            634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             357
                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                            ò
                                                                                                                           Q
                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                  env polyprotein - mouse mammary tumor virus
N:Contains: coat protein gp36; coat protein gp52
C:Species: mouse mammary tumor virus, MMTV
C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: 526388
R:Majors, J.E.; Varmus, H.E.
J. Virol. 47, 495-504, 1983
A;Title: Nucleotide sequencing of an apparent proviral copy of env mRNA defines determin A;Reference number: S26388; MUID:84011009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Cross_references: EMB.:K00556; NID:g332097; PIDN:AAA46533.1; PID:g332098 A;Note: the authors translated the codon TAT for residue 414 as Thr C;Genetics:
                       QINDLRQTVIWMG-DRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQG--R 470
                                                    -LGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  G------KGDKRRMWELWLTLGNSGA---NTKLVPIKKKLP--PKYPHCQIAFKKDA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240 FWEGDESAPPRWLPCAFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYT 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ESLDKHKHKKLQSFYL---WEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWS 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------KRPGFQEHEM-----IPTSACVTYPYVILLGLPQLIDIEKRGSTFHISCSSC 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RLFTCIDSTFNWQHRILLVRAREGMWIPVST-DRPW---EASPSIHILTEILKGVLNRSK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLAS 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414 QINDLRQTVIWMG-DRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQG--R 470
                                                                                                                                                   :: :: | | : | | : | | : | | : | | DABELSYNIQELTNLISDMSKQHIDAVDLSGLAQSFANGVKALNPLDWTQYF-IFIGVGAL 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113 PVGWGNTDPIRVLTNQTIYLGG-----SPDFHGFR-NMSGNVHFEEKSDTLPICFSF 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNC----- 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 PVTWMD-NPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYP-----FIC--- 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PPVNKEVHRWYEAGWVEPTWFWEN---SPKDPNDRDFTALVPHTELFRLVAASRYLIL--
                                                                                                                        EDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              243 GNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVG-----NIVIKPASQTITCENC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----AQVSPAVD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SFSTPTGCFQVDKQVFLSDTPTVDNN-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 12.9%; Score 388; DB 2; L
Best Local Similarity 23.4%; Pred. No. 2.2e-23;
Matches 144; Conservative 102; Mismatches 217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Gene: env
C;Superfamily: type A retrovirus env polyprotein
C;Keywords: glycoprotein; polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --SGQTQSCPS------
                                                                                                                                                                                                                                                      |:|: | :|
LLVIVLMIFPIVFQC 648
                                                                                                                                                                                                                       531 LIVVCLFCLLLVCRC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: DNA
A; Residues: 1-688 <MAJ>
                       414
                                                                      516
                                                                                                                                                                     575
                                                                                                                                                                                                                                                                         634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                354
                                                                                                                      471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                g
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                       ò
                                                                    g
                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
```

OSVNF SGTK- SDFCI II EALCI AIAGV	Qy 552 KTVT 555 1	Query Match 4.4%; Score 131; DB 2; Length 1420; Best Local Similarity 18.8%; Pred. No. 0.047; Pred. Similarity 18.8%; Pred. No. 0.047; Matches 112; Conservative 83; Mismatches 192; Indels 208; Gaps 32; Qy 70 VEVPTVSPNSRFTYHWSGMSLRPRVNLLODESYQRSLKFRPKGRTCPKEIPKGSK 125	Qy 181 DSD	Qy 345LKGVLNRSKRFIFTLIAVIMGLIAVTATAAN 375 Db 679 LSTEGQVQDFKFPLGIKGAGSSIQLSANTVKQNSRNGLAKLVFIIYRSLGQFLSTENATI 738 Qy 376AGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWM 425 1 1 1 1 1 1 1 1 1 1
QY 392 YWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITP 450 1 1 1 1 1 1 1 1 1 1	WESULT 6 WASIA IN polyprotein precursor - mouse intraci Alternate names: cost polyprotein Contains: surface protein; transmembran Species: mouse intracisternal A-particl Note: host Mus musculus (house mouse) Date: 31-Mar-1993 #sequence_revision 31 Accession: A41305 Reuss, F. U.; Schaller, H.C. Virol. 65, 5702-5709, 1991 Title: cDNA sequence and genomic charac Reference number: A41305; MUID:92015466 Molecule type: mRNA Residues: 1-584 < REU.> Residues: 1-584 < REU.> Cross-references: GB:M73818 Note: readthrough of three terminators Comment: This particle is a defective r Genetics:	C:Superfamily: intracisternal A-particle env polyprotein C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein F:1-25/Domain: extracellular #status predicted <ext> F:1-25/Domain: signal sequence #status predicted <ext> F:26-364/Product: env polyprotein #status predicted <env> F:26-362/Product: env polyprotein #status predicted ^{F:26-362/Region: urface processing #status predicted F:359-362/Region: cleavage processing #status predicted <imp> F:359-364/Product: transmembrane protein #status predicted <imp> F:367-354/Product: transmembrane protein #status predicted <imp> F:26-364/Product: transmembrane #status predicted <imi>F:364-302/Region: hydrophobic #status predicted <imi>F:548-584/Domain: intracellular #status predicted <imi>F:19,58,77,98,129,140,147,230,276,285,311,319,463,469,481,501/Binding site: carbohydrate</imi></imi></imi></imp></imp></imp>}</env></ext></ext>	Suery Match 5.0%; Score 152; DB 1; Length 584; Best Local Similarity 20.9%; Pred. No. 0.00026; Matches 101; Conservative 74; Mismatches 187; Indels 122; Ga 117 PKEIPKGSKNTEVLVWEECVANSVVILONNEFGTIIDXAPRGOF 157 PKIAPHCSLEDEGLILPWSDC-OSSITRWVDGSKTFSFSPMMIDD-PEKEFVMKKGLFIQ 161YHNCSGOTGSCPSAQVSPADSDLTESLDKHKKKLQSFYLWEWEEKGISTP 15	QY 265 TVPLQSCLRPPYMLVVGNIVIRPASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWI 324 1 1 1 1 1 1 1 1 1 1

S

ďį

```
Matches 112; Conservative
                                                                                                                                                                                                                                                                                                                                   A; Molecule type: mRNA
A; Residues: 1-1463 <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 DSD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                              Figures 1. Lellanova, V.G.; Ushkaryov, Y.A.

Figures 1. Lellanova, V.G.; Ushkaryov, Y.A.

Figures 1. Lellanova, V.G.; Ushkaryov, Y.A.

Figures 1. Lellanova, V.G.; Ushkaryov, Y.A.

Figures 1. Lellanova, 1999

Figures 1. Lellanova, 18869; MUID:99148828

Figures 1. Lellanorary, Translated from GB/EMBL/DDBJ

Figures 1. Lellanorary, Translated from GB/EMBL/DDBJ

Figures 1. Lellanorary, Translated from GB/EMBL/DDBJ

Figures 1. Lellanorary

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      32;
                                                                                                                                                                                                                                                                                                                          C;Species: Bos primigenius taurus (cattle)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T18387
R;Matsushita, H; Lelianova, V.G.; Ushkaryov, Y.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               425
                               822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  126 NTEVLVWEECVANSVVILQNNEF----GTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             679 LSTEGOVQDFKFPLGIKGAGSSIQLSANTVKQNSRNGLAKLVFIIYRSLGQFLSTENATI 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          739 KLGADFIGRNSTIAVNSHVISVSI----NKESSRVY--------LTDPVLFT 778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           70 VEVPTVSPNSRFTYHMVSGMSLRPRVNYL----QDFSYQRSLKFRPKGKTCPKEIPKGSK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VDVP--FPNQ---YQYIAAVDYNPRDNQLYVWNNNFILRYSLEFGPPD---PAQVP---- 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----TTAVTITSSAEMFKTTVSTTSTTSQKGPMSTTVAGSQEGSKGTKAPPAV 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                426 GDRLDLEHHFQLQCD-WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQ 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STTKIPPVTNIFPLPERFCEALDARGIR------WPQTQRGMMVERPCPKGTRGTASYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             508 CVLSTGTWNPKGPDLSNCT--SHWVNQLAQKIRSGENAASLANELAKHTKGPVFAGDVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -KIISPVSGPEHPELWRLTVASHHI-----RIWSGNQT-----LETRYRKPFYTIDLNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----ILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        566 SVRLMEQLVDILDAQLQE-LKPSEKDSAGRSYNKLOKREKTC-RAYLKAIVDTVDNLLRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----WQH-------RILLVRAREGMWI------PVSTDRPWEASPSIHILTEI--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----AGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LPHIDPDNYFNANCSFWNYSE-----RTMMGYWS-----TOGCK---LVDTNKTRTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 DSD------LTESLDKHKHKKLQSFYLWEWEEKGISTPRP------
                                                                                                                             823 CACSHLTNFAILMAHREIA-YKDGVHELLLTVITWV----GIVISLVCLAICIF 871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          485 IFEASKAHLNLVPGTEAIAGVADGLAN--LNPVTWIKTIRSTMIINLILIVVCLF 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CACSHLTNFAILMAHREIA-YKDGVHELLLTVITWV----GIVISLVCLAICIF 871
                                                                                       485 IFEASKAHLNLVPGTEAIAGVADGLAN--LNPVTWIKTIRSTMIINLILIVVCLF 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 4.4%; Score 131; DB 2; Length 1435; Best Local Similarity 18.8%; Pred. No. 0.047; Matches 112; Conservative 83; Mismatches 192; Indels 20
                                                                                                                                                                                                                                                                                                    latrophilin-2 (splice variant bbbbe) - bovine
779 LPHIDPDNYFNANCSFWNYSE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               263
                            8
                                                                                       ç
                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
```

σ

RESULT

```
larcophilin-2 (splice variant bbbbf) - bovine
C;Species: Bos primiqenius taurus (cattle)
C;Species: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T1838
R;MatSushita, H; Lelianova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A;Title: The latrophilin family: multiply spliced G protein-coupled receptors with d1
A;Reference number: 218869; MUID:99148828
                                                                                                                                                                                                                                           with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32;
latrophilin-2 (splice variant bbbaf) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Species: Bos primigenius taurus (cattle)
C;Species: 15-oct-1999 #sequence_revision 15-oct-1999 #text_change 18-Feb-2000
C;Accession: T1896
R;Matsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A;Title: The latrophilin family: multiply spliced G protein-coupled receptors A;Reference number: 218869; MUID:99148828
A;Accession: T18386
A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIDN: AAD05318.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               126 NTEVLVWEECVANSVVILQNNEF----GTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SVRLMEQLVDILDAQLQE-LKPSEKDSAGRSYNKLQKREKTC-RAYLKAIVDTVDNLLRP 623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----WQH-------RILLVRAREGMMI------PVSTDRPWEASPSIHILTEI-- 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----LKG------VLNRSKRFIFTLIAVI---MGLIAVTATAAV 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----AGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWM 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       70 VEVPTVSPNSRFTYHMVSGMSLRPRVNYL----QDFSYQRSLKFRPKGKTCPKEIPKGSK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LSTEGOVODFKFPLGIKGAGSSIQLSANTVKQNSRNGLAKLVFIIYRSLGQFLSTENATI 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GDRLDLEHHFQLQCD-WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQ 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |:| :| :| :| | EALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPTE----NIVLEVAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       508 CVLSTGTWNPKGPDLSNCT--SHWVNQLAQKIRSGENAASLANELAKHTKGPVFAGDVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KLGADFIGRNSTIAVNSHVISVSI----NKESSRVY-----LTDDVLFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      779 LPHIDPDNYFNANCSFWNYSE-----RTMMGYWS----TQGCK---LVDTNKTRTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----TTAVTITSSAEMFKTTVSTTSTTSQKGPMSTTVAGSQEGSKGTKAPPAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           455 STTKIPPVINIFPLPERFCEALDARGIR-----WPQTQRGMMVERPCPKGTRGTASYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -KIISPVSGPEHPELWRLTVASHHI-----RIWSGNQT-----LETRYRKPFYTIDLNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----ILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              823 CACSHLTMPAILMAHREIA-YKDGVHELLLTVITWV----GIVISLVCLAICIF 871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  485 IFEASKAHLNLVPGTEAIAGVADGLAN--LNPVTWIKTIRSTMIINLILIVVCLF 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   208;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Cross-references: EMBL:AF111082; NID:94164046; PID:94164047;
C;Superfamily: alpha-latrotoxin receptor, calcium-independent
C;Keywords: alternative splicing; G protein-coupled receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83; Mismatches 192; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----LTESLDKHKHKKLQSFYLWEWEEKGISTPRP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 0.049;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 4.4%; Score 131; Similarity 18.8%; Pred. No. 0
```

ဖ

```
protein-coupled receptors with di
                                                         32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1422 <AMT>
A;Residues: 1-1422 <AMT>
A;Cross-references: EMBL-AF111079; NID:g4164040; PID:g4164041; PIDN:AAD05315.1
C;Superfamily: alpha-latrotoxin receptor, calcium-independent
C;Keywords: alternative splicing; G protein-coupled receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      latrophilin-2, splice variant bbabe - bovine C;Species: Bos primigenius taurus (cattle) (C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000 C;Accession: T18383 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000 C;Accession: T18383 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000 FFBS Lett. 43, 348-352, 1999 A;Title: The latrophilin family: multiply spliced G protein-coupled receptor A;Accession: T18383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SVRLMEQLVDILDAQLQE-LKPSEKDSAGRSYNKAIVDTV--DNLLRPEALES---WKHM 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             126 NTEVLVWEECVANSVVILONNEF----GTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                NTEVLVWEECVANSVVILQNNEF-----GTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                              70 VEVPTVSPNSRFTYHMVSGMSLRPRVNYL----QDFSYQRSLKFRPKGKTCPKEIPKGSK 125
                                                                                                                                                                                                                                                                    : | : | : | 111 : | : | : | ---TTAVTITSSAEMFKTTVSTTSTTSQKGPMSTTVAGSQEGSKGTKAPPAV 454
                                                                                                                                                                                                                                                                                                                                                                                              STTKIPPVTNIFPLPERFCEALDARGIR-----WPQTQRGMMVERPCPKGTRGTASYL 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -AGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRLDLEHH 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FQLQCD-WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHL 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VEVPTVSPNSRFTYHMVSGMSLRPRVNYL----QDFSYQRSLKFRPKGKTCPKEIPKGSK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :: |: |: |: |: |: | CVLSTGTWNPKGPDLSNCT--SHWVNQLAQKIRSGENAASLANELAKHTKGPVFAGDVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | : | : | : | : | : | | SSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPTE-----NIVLEVAVLSTEGQVQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    775 FNANCSFWNYSE-----RTMMGYWS-----TQGCK---LVDTNKTRTTCACSHLTNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VDVP--FPNQ---YQYIAAVDYNPRDNQLYVWNNNFILRYSLEFGPPD---PAQVP---
                                                                                                                                                                                                                                                                                                                                          DSD------LTESLDKHKHKKLQSFYLWEWEEKGISTPRP-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                        -KIISPVSGPEHPELWRLTVASHHI-----RIWSGNQT-----LETRYRKPFYTIDLNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----RILLVRAREGMWI------PVSTDRPWEASPSIHILTEI------
Length 1407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1422;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----LKG-----VLNRSKRFIFTLIAVI---MGLIAVTATAAV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NLVPGTEAIAGVADGLAN -- LNPVTWIKTIRSTMIINLILIVVCLF 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4.2%; Score 125.5; DB 2;
18.8%; Pred. No. 0.13;
wtematches 189;
     DB 2;
                                                         189;
  Score 125.5; Di
Pred. No. 0.13;
                                                            Mismatches
                                                         84;
  4.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Simi
Matches 110;
                               Best Local Sim
Matches 110;
     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      358
                                                                                                                                                                     358
                                                                                                                                                                                                                               126
                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                              455
                                                                                                                                                                                                                                                                                                                                                                                                                                                        215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70
                                                                                                                                                                                                                                                                                   406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                  qq
                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                рp
                                                                                                                                                                                                                                                                                                                                                                                           D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ωý
                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C. Species: Bos primigenius taurus (cattle)
C. Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C. Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C. Accession: T18381
R. Matsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A.
R. Matsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A.
R. Title: The latrophilin family: multiply spliced G protein-coupled receptors with diffe A. Reference number: 218869; MUID: 99148828
A. Accession: T18381
A. Accession: T18381
A. Accession: T18381
A. Molecule type: mRNA
A. Residues: 1-1407 cMAT>
A. Residues: 1-1407 cMAT>
A. Residues: 1-1407 cMAT>
C. Superfamily: alpha-latrotoxin receptor, calcium-independent
C. Superfamily: alpha-latrotoxin receptor, calcium-independent
C. Superfamily: alpha-latrotoxin receptor.
C. Superfamily: alpha-latrotoxin receptor.
                                                                                                                                                                                                                                                                                                             32;
                                                                                                              PIDN: AAD05320.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----LKG------VLNRSKRFIFTLIAVI---MGLIAVTATAAV 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              679 LSTEGQVQDFKFPLGIKGAGSSIQLSANTVKQNSRNGLAKLVFIIYRSLGQFLSTENATI 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------AGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWM 425
                                                                                                                                                                                                                                                                                                                                                                 70 VEVPTVSPNSRFTYHMVSGMSLRPRVNYL----QDFSYQRSLKFRPKGKTCPKEIPKGSK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GDRLDLEHHFQLQCD-WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQ 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : : | : | : | : | : | : | : | : | CVLSTGTWNPKGPDLSNCT--SHWVNQLAQKIRSGENAASLANELAKHTKGPVFAGDVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KLGADFIGRNSTIAVNSHVISVSI----NKESSRVY------LTDPVLFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LPHIDPDDNYFNANCSFWNYSE-----RTMMGYWS-----TQGCK---LVDTNKTRTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               126 NTEVLVWEECVANSVVILQNNEF----GTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------TTAVTITSSAEMFKTTVSTTSTTSQKGPMSTTVAGSQEGSKGTKAPPAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STIKIPPVINIFPLPERFCEALDARGIR------WPQTQRGMMVERPCPKGIRGTASYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -KIISPVSGPEHPELWRLTVASHHI-----RIWSGNQT-----LETRYRKPFYTIDLNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SVRLMEQLVDILDAQLQE-LKPSEKDSAGRSYNKLQKREKTC-RAYLKAIVDTVDNLLRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----WQH-------RILLVRAREGMWI------PVSTDRPWEASPSIHILTEI--
                                                                                                                                                                                                                                                                                                                                                                                                                     358 VDVP--FPNQ---YQYIAAVDYNPRDNQLYVWNNNFILRYSLEFGPPD----PAQVP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----ILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           485 IFEASKAHLNLVPGTEAIAGVADGLAN--LNPVTWIKTIRSTMIINLILIVVCLF 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CACSHLTNFAILMAHREIA-YKDGVHELLLTVITWV----GIVISLVCLAICIF 871
                                                                                                                                                                                                                                                                                                             Mismatches 192; Indels 208;
                                                                                                                                                                                                                                                Length 1478;
A.Accession: T18388
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1478 <MAT>
A;Cross_references: EMBL:AF111084; NID:g4164050; PID:g4164051;
C;Superfamily: alpha-latrotoxin receptor, calcium-independent
C;Reywords: alternative splicing; G protein-coupled receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DSD-------LTESLDKHKHKKLQSFYLWEWEEKGISTPRP-
                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                   Query Match 4.4%; Score 131; DB 2 Best Local Similarity 18.8%; Pred. No. 0.049; Matches 112; Conservative 83; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        latrophilin-2 (splice variant bbaae) - bovine
```

215

οŽ

g

208 263

QΩ

406 181 455

요 ò g δ 999

a

ò

309

ŏ

624

QQ ò g

345

739 426

g ò 823

õ g

g

ĺ

376

ò

Db 508 CVLSTGTWNPKGPDLSNCTSHWVNQLAQKIRSGENAASLANELAKHTKGPVFAGDVSS 565 Qy 263ILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFNWOH- 311	RESULT 14 T18384 latrophilin-2 (splice variant bbabf) - bovine C;Specides: Bos primigenius taurus (cattle) C;Dactes: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000 C;Accession: T18384 R;Matsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A. FFRS Left 443 348-352 1999	A; Title: The latrophilin family: multiply spliced G protein-coupled receptors with di A; Reference number: 218869; MUD:99148828 A; Recession: T18384 A; Status: preliminary; translated from GB/EMBL/DDBJ A; Molecule type: mRNA A; Residues: 1-1465 - MANA A; Residues: 1-1465 - MANA A; Residues: 1-1465 - MANA A; Residues: EMBL:AF111080; NID:94164042; PID:94164043; PIDN:AAD05316.1 C; Superfamily: alpha-latrotoxin receptor, calcium-independent C; Superfamily: alpha-latrotoxin receptor, calcium-independent C; Keywords: alternative splicing; G protein-coupled receptor C; Keywords: alternative splicing; G protein-coupled receptor Best Local Similarity 18.8%; Pred. No. 0.14; Matches 110; Conservative 84; Mismatches 189; Indels 203; Gaps 32; Oy 70 VEVPTVSPNSRFTYHWYSGMSLRPRVNYLODFSYQRSLKFRPKGKTCPKEIPKGSK 125 Lill	Db 358 VDVPFPNQYQYIAAVDYNPRDNQLYVWNNNFILRYSLEFGPPDPAQVP 405 QY 126 NTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTGSCPSAQVSPAV 180 1
Db 406TTAVTITSSAEMEKTTVSTTSQKGPMSTTVGGSGEKGTKAPPAV 454 Qy 181 DSDTTAVTITSSAEMEKTTVSTTSTTSQKGPMSTTVGGSGEKGTKAPPAV 454 E	OY 435 FQLQCD-WNTSDFCITPQIYNESEHHWDMYRRHLQGREDNLTLDISKLKEQIFEASKAHL 493	RESULT 13 T18382 latrophilin-2 (splice variant bbaaf) - bovine C;Species: Bos primigentus taurus (cattle) C;Jpecies: Bos primigentus taurus (cattle) C;Jpecies: Bos primigentus taurus (cattle) C;Jpecies: Bos primigentus taurus (cattle) C;Accession: T18382 R;Matsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A. R;Matsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A. R;Matsushita: The latrophilin family: multiply spliced G protein-coupled receptors with diffe A;Reference number: Z18869; MUID:99148828 A;Reference number: Z18869; MUID:99148828 A;Retus: preliminary A;Rocusion: T18382 A;Molecule type: mRNA A;Cross-references: EMBL:AFII1078; NID:94164038; PID:94164039; PIDN:AAD05314.1 C;Superfamily: alpha-latrotoxin receptor, catclum independent	C; Keywords: alternative splicing; G protein-coupled receptor Query Match Best Local Similarity 18.8%; Pred. No. 0.13; Matches 110; Conservative 84; Mismatches 189; Indels 203; Gaps 32; Qy 70 VEVPTYSPNSPTTYHWYSGMSIRFRYNYLODESYQRSIKFRPKGKTCPKEIPKGSK 125 1: 1

Qy	345	LKGVLNRSKRFIFTLIAVIMGLIAVTATAAV 375
Qy Db	376 735	-AGVALHSSVOSVNEVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRLDLEHH 434 : : :: : : :: : :: : : : : : : :
Qy Db	435	FQLQCD-WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHL 493
Oy Dp	494	NLVPGTEAIAGVADGLANLNPVFWIKTIRSTMIINLILIVVCLF 537
RESULT T17158 CL2AB F C; Speci C; Date: C; Acces R; Sugit submitte A; Descr A; Refer	TESULT 15 TI7158 TI7158 TI7158 TI7158 C.Species: Ratt C.Species: Ratt C.Sate: 15-Oct C.Accession: TI R.Sugita, S.; 1 Submitted to th A.Description: A.R.Réference num	TITISB TITISB TITITSB TITITSB CL2AB protein - rat CL2AB protein - rat C;Species: Rattus norvegicus (Norway rat) C;Species: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000 C;Accession: TITISB R;Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C. submitted to the EMBL Data Library, July 1998 A;Description: CL family. A;Reference number: 218712
A; Ac A; Sta A; Mo A; Res A; Cr C; Su	A; Accession: A; Status: pre A; Molecule ty A; Residues: 1 A; Cross-refer C; Superfamily	A,Accession: T17158 A;Status: prellminary; translated from GB/EMBL/DDBJ A;Molecule type: mRNA A;Residues: 1.1420 <sug> A;Cross-references: EMBL:AF081149; NID:g3695124; PID:g3695125; PIDN:AAC62655.1 C;Superfamily: alpha-latrotoxin receptor, calcium-independent</sug>
Qu. Ma	Query Matc Best Local Matches 1	Match 3.9%; Score 118; DB 2; Length 1420; ocal Similarity 19.4%; Pred. No. 0.53; ss 115; Conservative 78; Mismatches 195; Indels 206; Gaps 33;
Qy Db	358	VEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSK 125 :
Qy Db	126	NTEVLVWEECVANSVVILQNNEFGTIIDXAPRGGFYHNCSGQTOSCPSAQVSPAVDSD 183
QY	184	LTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIIS 218 : :
Oy Db	219	PVSGPEHPELWRLTVASHHIRIWSGNOTLETRYRKPFYTIDLNS 262
Qy	263	30
Db	571	
Qy Db	309	WQH
Οy	345	LKGVLNRSKRFIFTLIAVIMGLIAVTATAAV 375
g	687	DFTFHLGFKGAFSSIQLSANTVKQNSRNGLAKVVFIIYRSLGPFLSTENATVKLGADLLG 746
Qy Db	376	AGVALHSSVQSVNFVNYWQKNSTRLANSQSSIDQKLASQINDLRQTVIWAGDRLDLEH 433 : :
ο,	434	HFQLQCD-WNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQI 485

787 YFNANCSFWNYSERIMMGYWSTQGCKLVDINKTRTTCACSHLIN 830	486 FEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLF 537	9
78.	48(83.
QQ	δλ	qq

Search completed: April 9, 2002, 16:58:14 Job time: 248 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

model
S
using
search,
protein
1
protein -
ŏ

Run on:

April 9, 2002, 17:07:01; Search time 23.68 Seconds (without alignments) 868.622 Million cell updates/sec

Title: Perfect score:

US-09-490-700-38 3011 1 MVTPVTWMDNPIEVYVNDSV......VCRCTPTAPKKTVTSRTGHE 561 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

100059 seqs, 36664827 residues Searched:

100059 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_39:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	uc	no sapien	mouse mamma		sheep pulmo		caprine art	bovine immu	no sapien					equine infe	simian immu	saccharomyc	1S Sarcom	bovine immu	æ	equine infe	equine infe	ovis aries	equine infe	feline leuk	rattus norv	s musculu	homo sapien	feline sarc	simian immu	ovine lenti	podospora a	variola vir	oryctolagus	xenopus lae
	ptic	omod /												_						~	_	-				Snw (_			_			_	7 xei
	Description	P10267	P10259	P03374	P3162	P31789	P3162	P1955	09548	P22429	P06751	P11306	P3254	P16082	P05885	P43596	P03396	P19557	P31626	P22427	P2243(P20757	P2242	P06752	P5259	062190	003001	P2144	· P11267	P1689	P2067	P3299	P3241	P3544
SUMMARIES	ID	ENV1_HUMAN	ENV_MMTVB	ENV_MMTVG	ENV_JSRV	ENV_IPMAE	ENV_CAEVG	ENV_BIV06	S24A_HUMAN	ENV_EIAV3	ENV_EIAVY	ENV_EIAV9	ENV_EIAVC	ENV_EIAVW	ENV_SIVM1	YFI3_YEAST	ENV_RSVP	ENV_BIV27	ENV_CAEVC	ENV_EIAV1	ENV_EIAV5	ANGT_SHEEP	ENV_EIAV2	ENV_FLVSA	N107_RAT	RON_MOUSE	BPA1_HUMAN	ENV_FSVSM	ENV_SIVML	ENV_OMVVS	NU5M_PODAN	VG05_VARV	FMO3_RABIT	FSPO_XENLA
	DB	-	Н	-	-		-	~	٦	٦	٦	٦	٦	٦	٦	-	Н	7	Н	-	-	-	-	-	~	-	1	Н	Н	Н	-	-	-	-
	Query Match Length	584	688	688	615	584	942	875	1078	859	859	829	859	859	882	787	603	904	996	859	860	476	859	639	926	1378	2704	645	880	066	652	3	530	803
æ	Query Match	90	13.2	٠. ش	11.3	٦.	3.9	٠,		3.6	3.6	•	•	3.5	3.5	3.5	3.4	3.4		3.4	3.4	3.3	3.3	3.3	3.3	3.5	3.5		3.5	3.5	3.1	3.1	3.1	3.1
	Score	2735	398	393	340.5	152	117	109.5	109	108	108	107	107	105	105	104	103.5	103.5	103	102	101.5	99.5	9	98.5	98	9	96.5	δ.	95	σ	94.5	94	94	93
	Result No.	-	7	e	4	S	9	7	6 0	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	30	31	32	33

LTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWSG 243

184

180

QQ οy

δ

124 121

Qy QQ 244 NQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCI 303

# 2 % # # # # # # # # # # # # # # # # #		.;
P03379 visna lenti P53753 saccharomyc P08359 feline leuk P32593 moloney mur P57740 homo sapien O9y210 homo sapien P1261 feline leuk P22380 simian immu P23422 visna lenti P12259 homo sapien	AA. (e) (ertebrata; Euteleostomi; Hominidae; Homo. H.; conac) (POTENTIAL).	Length 584; Indels 2; Gaps HYPPICLGRAPGCLMP 63 IIIIIIIIIIIIIIIII RYPPICLGRAPGCLMP 60 IRPPICLGRAPGCLMP 60 IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
ENV_VILV ENG_VENGL ENG_ENGENCEL KMOS_MSVMT NIO7_HUMAN ENV_RMCFV ENV_FLVLB ENV_FLVLB ENV_STVGB ENV_STVGB ENV_STVGB ENV_STVGB ENV_STVGB ENV_VILVI	ALIGNMENTS PRT; 584 AA. sequence update) sequence update) (PROTEIN. 121993; 121993; 17., Ushikubo H.; nan endogenous retrovirus genomerus genome."; POLYprotein. N-LINKED (GLCNAC) (POTEN N-LINK	SCOIE 2735; DB 1; Pred. No. 3.1e-219; 10; Mismatches 20; TDDRCPARPEEGMMINISIGS VSGMSLRPRVNYLQDFSYQRSI (VSGMSLRPRVNYLQDFSYQRSI VSGMSLRPRVNYLQDFSYQRSI
3.1 1982 1 3.1 1117 1 3.0 354 1 3.0 925 1 3.0 640 1 3.0 662 1 3.0 821 1 3.0 2224 1 3.0 5430 1	STANDARD; (Rel. 10, Creat (Rel. 10, Last (Rel. 10, Last ELATED ENV POL) (Human). etazoa; Chordat theria; Primate 606; M. N. A. M. N. A. Sequence of humanga T., Miyats 1589-598 (1986). VCHUER. 1. Coat protein, 1. Coat protein, 1. Coat protein, 257 257 257 257 257 257 257 257 257 257	90.8%; 1arity 94.1%; Conservative PIEVEVUNDSVWPGP
34 93 35 93 36 92.5 37 91.5 38 91.5 41 91.5 41 90.5 44 90.5	과 I	Ouery Match Best Local Similarity Matches 512; Conser 4 PVTWMDNPIEVYV 1
	RESULLA ENVILA LID AC LID DT DD DT DD DT OC OC OC OC OC OC OC OC OC OC	Ou Be Oo Oy Oy Oy

N

```
οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                          δ
                                                                                          d
                                                                                                                                        ŏ
                                                                                                                                                                           g
                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                   Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                        419
                                                                                                                                                                                                                                                        482
                                                                                                                                                                                                                                                                                                   479
                                                                                                                                                                                                                                                                                                                                                  542
                                                                                                                                                             MGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVI 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13.2%; Score 398; DB 1; Length 688;
llarity 23.7%; Pred. No. 3.3e-25;
Conservative 100; Mismatches 217; Indels 152; Gaps
DSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVI
                                                                                                                                                                                                                                                        WMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLK
                                                                                                                                                                                                                                                                                                                                                  EQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCLLLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  UL-1989 (Rel. 11, Last sequence update)
UG-2001 (Rel. 40, Last annotation update)
POLYPROTEIN (CONTAINS: COAT PROTEIN GP52; COAT PROTEIN GP36].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Moore R., Dixon M., Smith R., Peters G., Dickson C.; "Complete nuclectide sequence of a milk-transmitted mouse mamma tumor virus: two frameshift suppression events are required for translation of gag and pol."; J. Virol. 61:480-490(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (strain BR6).
Retroviridae; Betaretrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LINKED (GLCNAC. . .) (PA 46CAAC8D61FFADC2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. . . ) (1
N-LINKED (GLCNAC. . . ) (1
N-LINKED (GLCNAC. . . ) (1
N-LINKED (GLCNAC. . . ) (1
N-LINKED (GLCNAC. . . ) (1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COAT PROTEIN GP52.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        688 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LEADER PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-87112944; PubMed-3027377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-1989 (Rel. 10, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   77176 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M15122; AAA46544.1; -.
PIR; D26795; VCMVMM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mouse mammary tumor virus
Viruses; Retroid viruses;
VCBI_TaxID=11758;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     474
6688
1127
1143
2297
498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99
475
127
143
297
498
557
688 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                           CRCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CRCT 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUL-1989
20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ENV_MMTVB
P10259;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ENV_MATVB

AC POLOZÓS

DT DAZ POLOZÓS

DT OL-MAJ

DT OL-MAJ

DT OL-MAJ

DT OL-MAJ

DT OL-MAJ

DT OL-MAJ

OC VITUR

RA MEDLI

RA COR I

CC OR S

CC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                           480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      540
                       240
                                                                     304
                                                                                                             300
                                                                                                                                                           364
                                                                                                                                                                                                        360
                                                                                                                                                                                                                                                     424
                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                483
                                                                                                                                                                                                                                                                                                                                                                                                                                           543
                                                                                                                                                        õ
                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                  à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          env gene.";
--SPDFHGFR-NMSGNVHFEGKSDTLPICLSF 163
                                       --GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                 113 GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNC----- 164
                                                                                                                                                        G-----KGDKRRMWELWLTTLGNSGA---NTKLVPIKKKLP--PKYPHCQIAFKKDA 239
                                                                                                                                                                                                                                       240 FWEGDESAPPRWLPCAFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKNLFGNYT 299
                                                                                                                                                                                                                                                                              ESLDKHKHKKLQSFYL---WEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWS 242
                                                                                                                                                                                                                                                                                                                                                                                  RLFTCIDSTFNWQHRILLVRAREGMWIPVST-DRPW---EASPSIHILTEILKGVLNRSK 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                RFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLAS 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNNLHRNVTLALSEQRIIDLKLEA 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QINDLRQTVIWMG-DRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMYRRHLQG--R 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLI 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --SGQTQSCPS-----SDLT
                                                                                                                                                                                                                                                                                                           PPVNKEVHRWYEAGWVEPTWFWEN---SPKDPNDRDFTALVPHTELFRLVAASRHLIL--
                                                                                                                                                                                                                                                                                                                                                         243 GNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVG----NIVIKPASQTITCENC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GP36]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Redmond S.M.S., Dickson C.;
"Sequence and expression of the mouse mammary tumour virus EMBO J. 2:125-131(1983).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COAT PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mouse mammary tumor virus (strain GR).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUL-1989 (Rel. 11, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN [CONTAINS: COAT PROTEIN GP52;
                                                           688 AA
113 PVGWGSTDPIRVLTNOTMYLGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Rel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X01811; CAA25955.1; -. EMBL; K01788; AAA46537.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; A03972; VCMVM.
Coat protein; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |:|: | :| :|
634 LLVIVLMIFPIVFQC 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           531 LIVVCLFCLLLVCRC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11760;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-JUL-1986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ENV_MMTVG
P03374;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
ENV_MMTVG
                                                                                                                                                        193
                                                                                                                                                                                                165
                                                                                                                                                                                                                                                                            186
                                                                                                                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                                                                                                                                                                                                                       298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471
                                       55
                                                                                                                                                                                                                                                                                                                                                                                                  355
```

25;

54

Similarity

Query Match Best Local Simil Matches 146; (

ò

m

```
PubMed=1629959;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     M.W.
                                                                                                                                                                                                                                                                                                                      EMBL; M80216; AAA89184.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     69343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 23.8 tes 139; Conservative
                                                                                                                                                                                                                                                                                                                                                                                            80
379
80
379
403
108
127
178
275
319
319
   NCBI_TaxID=11746;
                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
TRANSMEM
DOMAIN
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   332
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                             25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   455
                                                                                                                                                                                                                                                                                                                   113 PVGWGSTDPIRVLTNQTMYLGG------SPDFHGFR-NMSGNVHFEGKSDTLPICFSF 163
                                                                                                                                                                                                                                                                                                                                                       54 -LGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 G------KGDKRRMWELWLHTLGNSGA---NTKLVPIKKKLP--PKYPHCQIAFKKDA 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLAS 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QINDLRQTVIWMG-DRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQG--R 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLI 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
20-JUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) (CONTAINS: COAT PROTEIN
                                                                                                                                                                                                                                                                                  PVTW-MDNPIEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGYHYP-----PIC--- 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --SGQTQSCPS-----SDLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVG-----NIVIKPASQTITCENC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RLFTCIDSTFNWQHRILLVRAREGMWIPVST-DRPW---EASPSIHILTEILKGVLNRSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ESLDKHKKKKLQSFYL---WEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RFVAAIILGISALIAIITSFAVATTALVKEMOTATFVNNLHRNVTLALSEQRIIDLKLEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FWEGDESAPPRWLPCAFPDKGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------KRPGFQEHEM-----IPTSACVTYPYAILLGLPQLIDIEKRGSTFHISCSSC
                                                                                                                                                                                                                                                                                                                                                                                                                               GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sheep pulmonary adenomatosis virus (Jaagsiekte sheep retrovirus)
                                                         (POTENTIAL)
                                                                                                                                                                                                                                           Indels 152;
                                                                                                                                                                                                      Query Match 13.1%; Score 393; DB 1; Length 688; Best Local Similarity 23.6%; Pred. No. 8.5e-25; Matches 145; Conservative 99; Mismatches 219; Indels 1
               COAT PROTEIN GP52.
COAT PROTEIN GP36.
N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . . ) (POT N-LINKED (GLCNAC. . ) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (JSRV).
Viruses; Retroid viruses; Retroviridae; Betaretrovirus
                                                                                                                                                                                                                                                                                                                                                                                          SFSTPTGCFQVDKQVFLSDTPTVDNN------
   LEADER PEPTIDE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              615 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                  Ψ.,
                                                                                                                                                  77220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GP52; COAT PROTEIN GP36].
98
474
688
1127
1143
297
557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |:|: | :|
634 LLVIVLMIFPIVFQC 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LIVVCLFCLLLVCRC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                688 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENV_JSRV
P31621;
                                                        CARBOHYD
CARBOHYD
                                                                                                            CARBOHYD
CARBOHYD
                                                                                            CARBOHYD
                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                          164
                                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 'ENV_JSRV
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AC DE DE OS OS OS OS
 FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TCPKEIPKGSKNTEVLV----WEECVANSVVILQ-NNEFGTIIDXAPRGQFYHNCSGQT 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          298 YPFMLIQGHMEITLSLNIYHLNCSNCILTNCIRGVAKGE-QVIIVKQPAFVMLPVEIAEA 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                      exogenous and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       451 QIYNESEHHWDMVRRHLQG--REDNLTLDISKLKEQIFE---ASKAHLNLVPGTEAIAGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QSCPSAQVSPAVDSDLTESLDKHKKKLQSFYLWEWEEKGISTP-----RPKIISPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----TPDEPEYRGOHPPIFSVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WYDETALELLORI-NTALSRPKRGLSLIILGIVSLITLIATAVTACVSLAQSIQAAHTVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          392 YWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           145 SQHPHCIQVSAD---ISYPRVT----ISGID------EKTGKKSYGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      180 TGPLDIPFCDKHLSIGIGIDTPWTLCRARVASVYNINNANAT-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              221 SGPEH-PELWRLTVASHHIRIWSGNQTLETRYRKP-----FYTIDLNSILTVPLQSCLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         274 PPYMLVVGNIVIKPASQT -- ITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINIS---IGYHYPPICL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11.3%; Score 340.5; DB 1; Length 615; 23.8%; Pred. No. 1.7e-20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COAT PROTEIN GP52 (POTENTIAL)
COAT PROTEIN GP36 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CYTOPLASMIC (POTENTIAL).

N-LINKED (GLCNAC...) (POT

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EXTRACELLULAR (POTENTIAL). POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIR; E42740; VCMVJA.

Coat protein; Glycoprotein; Polyprotein; Transmembrane.
PROPEP 1 79
York D.F., Vigne R., Verword D.W., Querat G.; "Nucleotide sequence of the jaagsiekte retrovirus, an endogenous type D and B retrovirus of sheep and goats. J. Virol. 66:4930-4939(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100; Mismatches
```

```
265 TVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWI 324
                                                                                                                                                                                                                                                                                                                                                                                SIN STATE ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSV 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PVPVETPSTLS-------MFRQKRDFGITAAMIIAISASAAAATAAGYAMVSAV 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRLDLEHHF-QLQCDWNT 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGTK-LNQLSADLADAITVQTSASTKLKGGLMILNQCLDLAEEQIGVLHQMAQLGCERKL 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SDFCITPQIYNESEHHWDMVRR---HLOGREDNLTLDISKLKEQIFEASKAHLNLVPGTE 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      213 RPKIISPVSGPEHPELWRLTVASH-----HIRIWSGNQTLETRYRKPFYTIDLNSIL 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          501 AIAGVADGLANLNPVTWIKT----IRSTMIINLILIVVC---LFCLLLVC--RCTPTAPK 551
                                                                                                                                                    DFRMHPFHKWVLCGVNGSC - - TELNPLI - - - - - - - - FIQGGAVGKASFTGIS - -
                                                                                       ----YHN--CSGQTQSCPSAQVSPAVDSDLTESLDKHKKKLQSFYLWEWEEKGISTP
                                                                                                                                                                                                                                                                               -----RFAQYWGIHDASQDSYGYTNTSVEITGFNKTLVNQINYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     552 KTVT 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               560 VVIT 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
ENV_CAEVG
                                                                                                                                                    215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      444
                         157
                                                                                          161
                                                                                                                                                                                                                                                                                  257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                       δ
                                                                                                                                                    g
                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mouse intracisternal a-particle (IAP-MIAE).
Viruses; Retroid viruses; Retroviridae; Intracisternal A-particles
--IADT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MISCELLANBOUS: READTHROUGH OF THREE TERMINATORS OCCURS: UGA BETWEEN CODONS FOR 71-THR AND 72-ALA, UGA BETWEEN CODONS FOR 111-HIS AND 112-ARG, AND UAA BETWEEN CODONS FOR 394-VAL AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                         01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
20-JUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: SURFACE PROTEIN; TRANSMEMBRANE PROTEIN].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Transmembrane; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-92015460; PubMed-1920613;
Reuss F.U., Schaller H.C.;
"cohx sequence and genomic characterization of intracisternal
A-particle-related retroviral elements containing an envelope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MISCELLANEOUS: THIS PARTICLE IS A DEFECTIVE RETROVIRUS.
476 KPYNTSDFPWDKVKKHLQGIWFNTNLSLDLLQLHNEILDIENSPKATLN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     180CCAF8F14A4CD6 CRC64;
                                                             ADG-LANL-NPVTWIKTIRSTMI --- INLILIVVCLFCLLLVCR 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEMBRANE PROTEIN. POTENTIAL.
                                                                                             529 VDNFLONLFSNFPSLHSLWKTLIGLGIFVIIIAIVIFVFPCVVR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ENV POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SURFACE PROTEIN
                                                                                                                                                                                                                                                                                  584 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M73818; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              J. Virol. 65:5702-5709(1991)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A41305; VCMSIA.
Coat protein; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65024
                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             311
319
463
469
481
501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  501
584 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26
26
363
363
527
58
77
77
98
129
140
                                                                                                                                                                                                                                                                            ENV_IPMAE
P31789;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                 RESULT 5
ENV_IPMAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                  HIDDON NEW RESET STATEMENT OF S
a
                                                             ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                         "Structure and genetic variability of envelope glycoproteins of two antigenic variants of caprine arthritis-encephalitis lentlvirus."; J. Virol. 65:5744-5750(1991).
                                                 Last sequence update)
Last annotation update)
RSOR (COAT POLYPROTEIN) [CONTAINS: SURFACE
                                                                                                                                                                                                         SEQUENCE FROM N.A., AND SEQUENCE OF 81-95.
MEDLINE-92015464; PubMed-1656067;
Knowles D.P. Jr., Cheevers W.P., McGuire T.C., Brassfield A.L.,
                                                                                                                                          G63) (CAEV)
                                                                                                                                                                                                                                                                                                                                                                      Knowles D.P.; Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases
                                                                                                                                        Caprine arthritis encephalitis virus (strain G63) Viruses; Retroid viruses; Retroviridae; Lentivirus NCBI_TaxID=11662;
   942 AA
   PRT;
                                  26, Created)
40, Last sequ
40, Last anno
                                                                                   ENV POLYPROTÈIN PRECURSOR (COAT
PROTEIN; TRANSMEMBRANE PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M60855; AAB88709.2; -.
PIR; A41307; VCLJC6.
InterPro; IPR000328; Env_GP41.
   STANDARD;
                                                                                                                                                                                                                                                                Stem T.A.;
                                                                                                                                                                                                                                                                Harwood W.G.,
                                                  20-AUG-2001
20-AUG-2001
ENV_CAEVG
P31627;
                                                                                                                                                                                                                                                                                                                                                    REVISIONS
```

24;

Gaps

5.0%; Score 152; DB 1; Length 584; 20.9%; Pred. No. 6.8e-05; .ive 74; Mismatches 187; Indels 122;

117 PKEIPKGSKNTE--VLVWEECVANSVV--ILQNNEFG---TIIDXAPRGQF----

101; Conservative

Similarity

Query Match Best Local

Matches

ò

550

-----TPTAP

| : : | ---EQVRRIPDV 787

S

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-91251255; PubMed-1445801; Oberste M.S., Greenwood J.D., Gonda M.A.; Analysis of the transcription pattern and mapping of the putative rev and env splice junctions of bovine immunodeficiency-like virus."; J. Virol. 65:3932-3937(1991).
                                                                          788 WESLKEVFDWSGWFSWLKYIP----IIVVGLVGCILIRAVICVCOPLVQIYRTLSTPTYQ 843
                                                                                                                                                                                                                                             01-FEB-1991 (Rel. 17, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                     Garvey K.J., Oberste M.S., Elser J.E., Braun M.J., Gonda M.A.; "Nucleotide sequence and genome organization of biologically active proviruses of the bovine immunodeficiency-like virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SURFACE PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Coat protein; Polyprotein; Transmembrane; Signal.
 738 INWTRFKDNCTWQQWERELQGYDGNLTM---LLRESARQTQLAE---
                                                                                                                                                                                                                                                                                                                                           Bovine immunodeficiency virus (isolate 106) (BIV). Viruses; Retroid viruses; Retroviridae; Lentivirus.
                                              LANLNPV-----TWIKTIRSTMIINLILIVVCLFCLLLVCRC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COAT PROTEIN GP62
COAT PROTEIN GP40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                 875 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLYCOPROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M32691; -; NOT_ANNOTATED_CDS
EMBL; M74711; AAA42762.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=90223985; PubMed=2183467;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virology 175:391-409(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-99 FROM N.A.
                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                              GP62; COAT PROTEIN GP40]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HIV; M32691; ENV$BIV106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             555
875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            722
131
2248
2267
3300
3300
403
403
462
515
634
634
                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11658;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           131
248
300
338
347
356
381
398
4403
4403
4403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glycoprotein;
SIGNAL
                                                                                                                                       844 RVTV 847
                                                                                                        551 KKTV 554
                                                                                                                                                                                                                 ENV_BIV06
P19556;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                             509
                                                                                                                                                                                                 ENV_BIV06
                                                                          g
                                                                                                       δ
                                                                                                                                      qq
                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             475 VIICRNNNTITGIMGTNTNWTTMWRIYRNCSG----CENATLDRKETGTLGGVANKNCSL 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    384 IKIGQETFSYHYKPNWNCTG------NWTQHPVWQVMRDLDMVEHMTGECVQRPQ 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RVNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEEC-----VANS---- 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RHNITVD------RNQTI-----TGNCSVTNWDGCNCSRSGNYLYNSTTGGLL 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VVILQNNE----FGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDK---- 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LTVASHHIRIWSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQ 290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : :: | :| : | : | : | : :: | EATYAMVQHVAKGVRILEARVARVEAITDRIML--YQELDC-WHYHQYCVTSTRADVAKY 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------HWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADG 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISIG----YHYPP--ICLGRAPGCLMPAVQNWLVE-VPTVSPNSRFTYHMVSGMSLRP- 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----VIKVRAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -HKHK------KLQSFYL-----WEWEEKGISTPRPKIISPVSGPEHPELWR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PHKNESNKWTCAPRQRECKTDSLYIAGGKKFWTREKAQYSCENN--IGELDGMLH----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITCENCRLFTCIDSTFNW-QHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | : |: |: || || || || T------YGVIEMPENYAKTRINRRKRE------LSH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NRSKRFI-FTLIAVIMGLIAVTATAAVAGVA-----LHSSVQSVNFVNYWQKNSTRLW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NSQSSIDQKLASQINDLRQTVI - - - WMGDRLDLEHHFQLQCDWNTSDFCITPQIYNESEH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                           POTENTIAL)
                                                                                                                                                      POTENTIAL)
                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                             POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                           TRANSMEMBRANE PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 1; Length 942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 117; DB 1; Length 94
Pred. No. 0.1;
1; Mismatches 178; Indels
                             LEADER PEPTIDE.
SURFACE PROTEIN (POTENTIAL)
                                                                                                  CYTOPLASMIC (POTENTIAL).

POTENTIAL.

N-LINEED (GLCNAC...) (POT N-LINKED (GLCNAC...)) (POT N-LINKED (GLCNAC...) (POT N-LINKED (GLCNAC...)) (POT N-LINKED (GLCNAC...) (POT N-LINKED (GLCNAC...)) (POT 
                Polyprotein; Transmembrane.
                                                                          EXTRACELLULAR (POTENTIAL)
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5B0E1F0F3D355F4A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----OGILLQKYQ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ,11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3.9%;
                Coat protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 108437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity 20.4
                                                                                                                                                                                                                                                                                          379
100
104
135
Pfam; PF00517; GP41; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ¥,
                                                                                                                                                                                 1131
1176
2228
3331
3348
3370
3370
4400
4435
                                                                                                                                                                                                                                                                                                                                                        441
447
457
467
481
503
              3lycoprotein;
PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Sim
hes 123;
                                                                                                     DOMAIN
TRANSMEM
DOMAIN
                                                                        DOMAIN
TRANSMEM
                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                    CARBOHYD
                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        584
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

(POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL)

(POTENTIAL)

(POTENTIAL).
(POTENTIAL).
(POTENTIAL).

(POTENTIAL)

(POTENTIAL) (POTENTIAL) (POTENTIAL)

(POTENTIAL)

```
34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-AUG-2001 (Rel. 40, Created)
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
PROTEIN TRANSPORT PROTEIN SEC24A (SEC24-RELATED PROTEIN A) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                          292
                                                                                                                                                                                                                                                                                                                                                                                                                                 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           645
                                                                                                                                                                                                                                                                                 382
                                                                                                                                                                                                                                                                                                                                                                                                      SLOFCVOWPYVLVGLNNSDSNISFNSGDWIAT----NCMHP-----ITLNKSAQDLG 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                         431 KNFPRLTFLDGQLSQLKNTLCGHNTNCLKFGNKSLSTNSLILCQDNPIGNDTFYSLSHSF 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVT----WIKTIRS 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QDEYDKIEEKILKIRVDWLNSSLSDTQDTFGLE--TSIFDHLVQLFDWTSWKDWIKII-- 701
                                                                                                                                                    WLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKG-KTCPKEIPKGSKN 126
                                                                                                                                                                              WLSANQEIPPKFAFPIILISS----BLAGIIGYYIMERHLEIFKKGCOIC-----GSLS 206
                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                      I-----SPVSGPEHPELWRLTVASHHIRIWSGNQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----RILLVRARE-GMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLIAVIM
                                                                                                                                                                                                                                                                                                                                             LEEBAVY ELALLSANDSRQVVVENGTDVCSSQNSSTNKGHP-MTLLKLRGQVSETWIGNS
                                                                                                                                                                                                                                                                                                                                                                            TLETRYRKPFYTIDLNS-----ILTVPLQSCLKPPYMLVVGNIVIKPASQTI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SRQASARWILVKVPSYGFVVVNDTDTP----PSLRIRKPRAVGLA-----IFLLVLAIM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A-ITSSLVAATTLVNQHTTAKVVERVVQNVSYIAQTQDQFTHLFRNINNRLNVLHRRVSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --QINDLRQTVIWMGDRLDLEH-HFQL---QCDWNTSDFCITPQIYNESEHHWDMVRRHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---LMWGMLLEEIGRRLARREWSVSRVVVILLISFSWGMYVNRVNASGSHVAMVTSPPGY
                                                                                                                                                                                                                                                                 -IIDXAPRGQFYHNCSGQTQSCPSAQVSPA-VDSDLTESLDKHKHKKLQSFYLWEWEEKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            365 GLIAVTATAAVAGVALHSS-----VQSVNFVNYWQKNSTRLW---NSQSSIDQKLAS-
                                                                                                                                                                                                                                                                                                                                                                                                                                 ----ENC---ENC------ENMOH--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
. .) (POTENTIAL)
                                                                                                                          199;
                                                                                               DB 1; Length 875;
                                                                                                                                                                                                         TEVLVW-----QNNEFGT----
                                                                                                                        Indels
N-LINKED (GLCNAC. . .) (POT
N-LINKED (GLCNAC. . .) (POT
K -> E (IN REF. 2).
K -> E (IN REF. 2).
W; EBEGFCD32747EA6C CRC64;
                                                                                            3.6%; Score 109.5; DB 1;
llarity 19.4%; Pred. No. 0.39;
Conservative 86; Mismatches 215;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1078 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE=B-cell;
MEDLINE=99175155; PubMed=10075675;
 665
848 N.
24 K
92 K
99166 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             524 TMIINLILIVVCLFCLLLVC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       702 IVIIVLWLLIKILLGMLRSC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human)
 665
848
24
92
875 AA;
                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=9606;
                                                                                            Query Match
Best Local Simi]
Matches 120; C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S24A_HUMAN
095486;
 CARBOHYD
CARBOHYD
                                        CONFLICT
                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEC24A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
S24A_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                646 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                491
                                                                                                                                                                                                                                                                                           264
                                                                                                                                                                                                                                                                                                                     509
                                                                                                                                                                                                                                                                                                                                                                             246
                                                                                                                                                                                                                                                                                                                                                                                                       383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      468
                                                                                                                                                    68
                                                                                                                                                                               157
                                                                                                                                                                                                          127
                                                                                                                                                                                                                                      207
                                                                                                                                                                                                                                                                 151
                                                                                                                                                                                                                                                                                                                                                324
                                                                                                                                                                                                                                                                                                                                                                                                                                 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               414
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FFF
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31;
                 OCT L., Paccaud J.-P.;
"Sec24 proteins and sorting at the endoplasmic reticulum.";
"Sec24 proteins and sorting at the endoplasmic reticulum.";
"Sec24 proteins and sorting at the endoplasmic reticulum.";
"Sec24 proteins and sorting at the COPIL Char. THAT COVERS ER-DERIVED
'-- FUNCTION: COMPONENT OF THE COPIL COAT, THAT COVERS ER-DERIVED
VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO
THE GOLGI APPRARTIUS. COPIL ACTS IN THE CYTOPLASM TO PROMOTE THE
TRANSPORT OF SECRETORY. PLASMA MEMBRANE, AND VACUOLAR PROTEINS
--- SUBUNIT: COPIL IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
COMPLEX, THE SEC13/31 COMPLEX AND SARI.
--- SUBGELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR (BY SIMILARITY).
--- TISSUE SPECIFICITY: EXPRESSED IN FIBROBLASTS, HEPATOCYTES, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      272 SVGYSYP-----SLPPGYQNTTPPGATGVPPSSLNYPSGPQAFTQTPLGANHLITS 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----SLGC---ISRYSAGSVYYYPSYHH------OHNPVQVQKLQKEL---Q 721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----HM---- 85
                                                                                                                                                                                                                                                                                 LYMPHOCYTES.
SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MSGLSLQPEGLRVVNLLQERNMLPSTPLKPPVPNLHEDIQKLNCNPELF---RCTLTSIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               188 LDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEH--PELWRLTVASHHIRIWSGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | ::|| | || || KELVQDILKTLPQMFTKTLETQSALGPAFKLMSPTGGRMSVFQTQTQTPTLGVGALKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86 VSGMSLRPR----VNYLQDFSYQRSLKFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TLETRYRKPFYTIDLNSILTVPLQSCLKPP----YMLV--VGNIVIKPASQTITCE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       472 TIE------FMAPSEYMLRPPQPPVYLFVFDVSHNAVETGYLNSVCQSLLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----OHRITLVRAREGMWIPVSTDRFWW------QHRITLVRAREGMWIPVSTDRPWEASPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38 IHILTEILK------GVLNRSKRFIFTLIAVIMGLIAVTAT-AAVAGVAL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----HSSVQSVNF---VNYWQKNSTRLWNSQSSIDQKLAS-QINDLRQTVIWMGDRLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 ILQNNEFGTIIDXA--PRGQFYHNCSGQTQ-----SCPSAQ--VSPAVDSDLTES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NLDLLPGNTRTKIGFITFDSTIHFYGLQESLSQPQMLIVSDIEDVFIPMPENLLVNLNES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REEPNHRSSAKDIHMTPSTDFYKKLALDCSGQQVAVDLFLLSGQYSDLA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AJ131244; CAA10334.1; -.
Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
Letourneur F., Garcia-Estefania D., Carpentier J.-L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1078;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             491 AHLNLVPGTEAIAGV--ADGLANLNPVTWIKT-----IRSTMIINL 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9A3214F52D28FD3F CRC64,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0.57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44 SIGYHYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTY--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZINC FINGER-LIKE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 0.57
; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1078 AA; 118791 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 3.6%; Sco
Best Local Similarity 19.8%; Pro
Matches 116; Conservative 85;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              323 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            431
      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
```

ò

^

셤

```
3.6%; Score 108; DB 1; Length 859;
21.0%; Pred. No. 0.51;
ve 83; Mismatches 237; Indels 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Equine infectious anemia virus (isolate Wyoming) (EIAV).
Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBL_TaxID=11672;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               859 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                           21.0%;
                           Best Local Similarity 21.0
Matches 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GP90; COAT PROTEIN GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENV_EIAVY
P06751:
  Query Match
Best Local 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ssel C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           515
                                                                                                                                                                                                                                                                167
                                                                                                                                                                                                                                                                                                                    146
                                                                                                                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                                                                                                      200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENV_EIAVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                          ò
                                                                                                                                                        qq
                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                      QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                     01-AUG-1991 (Rel. 19, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Payne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J., Montelaro R.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes reported for sequential isolates of HIV."; virology 161:321-331(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal SIGNAL
  | | :::| :::|
| THIPHGNFFVRSTDLLSL 761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    982A9F5A1AD8FA4D CRC64;
                                                                                                                                                                                                                                                                                                                                                       Equine infectious anemia virus (clone P3.2-3) (EIAV). Viruses; Retroid viruses; Retroviridae; Lentivirus. NCBI_TaxID=11668;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC, N-LINKED) (GLCNAC, N-LINKED (GLCNAC, N-LINKED) (GLCNAC, N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN GP90.
PROTEIN GP45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ENV POLYPROTEIN.
                                                                                                                                                           Ϋ́
                                                                                                                                                           859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COAT PROTEI
COAT PROTEI
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. MEDLINE-88072070; PubMed=2825406;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; C34027; VCLJE3.
InterPro; IPR001027; EIAV_GP45.
InterPro; IPR001361; EIAV_GP90.
Pfam; PF01045; EIAV_GP45; I.
Pfam; PF00971; EIAV_GP95; I.
                                                                                                                                                                                                            (Rel. 19, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97066 MW;
  EMBL; M18387; AAA66409.1; -
                                                                                                                                                                                                                                                                                                                    GP90; COAT PROTEIN GP45].
                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                557
752
859 AA;
                                                                                                                                                                                                            01-AUG-1991
                                                                                                                                                   ENV_EIAV3
P22429;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                             ENV_EIAV3
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    330 VPIFYTCTFTNI-----TSCNNEPIISVIMYETNQV-----QYLLCNNNSNNNNCVVQS 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lentivirus genomic organization: the complete nucleotide sequence of
                                                                                                                                                          113 GK-----TCPK-EIPKGSKNTEVLV-------WEECVAN---SVVILQN 145
                                                                                                                                                                                                                                                                                                                                                                                                            402 NSQSSIDQK---LASQINDLRQTVIWMGDRLDLEHHFQL-QCDWNTSDFCITPQIYNESE 457
                                                  53 CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                     122 CWGSFPGC--RPFQNY----FSYETNRSMHMDNNTATLLE-AYHREIAFIYK 166
                                                                                                                                                                                                               SSCTDSDHCQEYQCKKVNLNSSDSSNSVRVEDVTNTAEYWGFKWLECNQTEHFKTILVPE 226
                                                                                                                                                                                                                                                                     NEFGTIIDX---APRGQFYHNCS---GQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSF 199
                                                                                                                                                                                                                                                                                                                                                                          YLWEWEEKGIS-TPR----PKIISPVSGPEHPELWRLTVASHHIRIWSGNQTLETR-YR 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KP-FYTIDLNSILTVPLQSCLKPPYMLVV---GNIVIKPASQTITC--ENCRLFTCIDST 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FN--WQHRILLVRAREGM-----------WIPVSTDRPWEASPSIHILTEI 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LKGVLNRSKRF-IFTLIAVIMGLIAVTATAAVAGVALH--SSVQSVNFVNYWQKNSTRLW 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       458 HHWDMVRRHL--QGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNP- 514
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JAN-1988 (Rel. 06, Created)
01-JAN-1988 (Rel. 06, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  494 NGMDLIEROIKILYAMILQTHADVQLLKERQOVEETFNLIGCIERTHVFCHT-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             547 HPWNMSWGHLNESTQWDDWVSKMEDLNQEILTTLHGARNNLAQSMITFNTPDSIAQFGKD
                                                                                                                                                                                                                                                                                                      ----VTWIKTIRSTMIINLILIVVCLFCLLLVCRCTPTAPK----KTVTSRTG 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 607 LWSHIGNWIPGLGAS-IIKYIVMFLLIYLLL-----TSSPKILRALWKVTSGAG 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-87071653; PubMed-2431539;
Rushlow K., Olsen K., Stiegler G., Payne S.L., Montelaro R.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the env gene region of equine infectious anemia virus.";
Virology 155:309-321(1986).
```

```
CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                          Tronick
                                                                                             RESULT 11
ENV_EIAV9
                  458
                                  547
                                                  515
                                                                                                                     g
                                                                   g
                òγ
                                                   à
                                                                                                                                                                                                                                                                                                                                                  35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                273 FL--VQEKGIADISRIGNCGPIFLGVL-EDNKGVVRGDYTACNVRRLNINRKDYTGIYQ 329
                                                                                                                                                                                                                                                                                                                                                                                                  GK-----TCPK-EIPKGSKNTEVLV------WEECVAN---SVVILQN 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     200 YLWEWEEKGIS-TPR----PKIISPVSGPEHPELWRLTVASHHIRIWSGNQTLETR-YR 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      253 KP-FYTIDLNSILTVPLQSCLKPPYMLVV---GNIVIKPASQTITC--ENCRLFTCIDST 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VPIFYTCTFTNI-----TSCNSKPIISVIMYETNQV-----QYLLCNNNSNNYNCVVQS 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                  122 CWGSFPGC--RPFQNY------FSYETNRSMHMDNNTATLLE-AYHREITFIYK 166
                                                                                                                                                                                                                                                                                                                                                                                                                   167 SSCTDSDHCQEYQCKKVNLNSSASSNSVRVEDVTNTAEYWGFKWLECNQTENFKTILVPE 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FN--WQHRILLVRAREGM--------WIPVSTDRPWEASPSIHILTEI 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FGVIGOAHLELPRPNKRIRNOSFNOYNCSINNKTELETWKLVKT----SGVTPLPISSEA 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        345 LKGVLNRSKRF-IFTLIAVIMGLIAVTATAAVAGVALH--SSVQSVNFVNYWQKNSTRLW 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NSQSSIDQK----LASQINDLRQTVIWMGDRLDLEHHFQL-QCDWNTSDFCITPQIYNESE 457
                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                              NEFGTIIDX---APRGQFYHNCS---GQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSF
                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                       (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                3.6%; Score 108; DB 1; Length 859;
21.0%; Pred. No. 0.51;
tive 83; Mismatches 237; Indels 150;
                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                  Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal
SIGNAL
1 22
POTENTIAL.
                                                                                                                                                                                                                                                                                                         BC71EBD00FC5A2B9 CRC64;
                                                                                          COAT PROTEIN GP90.
                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                GLCNAC.
                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                      (GLCNAC.
                                                                                                                                                                              GLCNAC
                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                      GLCNAC.
                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                  POLYPROTEIN.
send an email to license@isb-sib.ch).
                                                                                                  COAT PROTE POTENTIAL. POTENTIAL.
                                                                                                                           POTENTIAL.
                                                                                                                                                             N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                     POTENTIAL
                                                                                                                                             POTENTIAL
                                                                                                                                                      N-LINKED
               EMBL; M16575; AAB59863.1; ALT_SEQ
                       PIR; A25610; VCLJEV.
InterPro; IPR001027; EIAV_GP45.
InterPro; IPR001361; EIAV_GP90.
Pfam; PF01045; EIAV_GP45; I.
Pfam; PF00971; EIAV_GP95; I.
                                                                                                                                                                                                                                                                                                         97041 MW;
                                                                                                                                                                                                                                                                                                                                        Best_Local Similarity 21.0%
Matches 125; Conservative
                                                                                  859
444
859
859
472
636
                                                                                                                                                                      752
859 AA;
                                                                                                                                                                      1141
1148
1186
2214
2233
340
340
440
4411
4483
550
550
                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                            TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                           TRANSMEM
                                                                                                                     TRANSMEM
                                                                                                                                              TRANSMEM
                                                                                                                                                      CARBOHYD
                                                                                                                                                             CARBOHYD
                                                                                                                                                                    CARBOHYD
                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                   CHAIN
                                                                                            CHAIN
                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                    227
                                                                                                                                                                                                                                                                                                                                                                  23
                                                                                                                                                                                                                                                                                                                                                                                                   113
                                                                                                                                                                                                                                                                                                                                                                                                                                    146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        402
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ·γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
494 NGMDLIERQIKILYAMILQTHADVQLLKERQQVEETFNLIGCIERTHVFCHT-----G 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-FEB-1996 (Rel. 33, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) (CONTAINS: COAT PROTEIN
                                                                                                                             HENNINSWGHLNESTQWDDWVSKMEDLNQEILTTLHGARNNLAQSMITFNTFDSIAQFGKD
                                                                        HHWDMVRRHL--QGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNP-
                                                                                                                                                                                                                    ----VTWIKTIRSTMIINLILIVVCLFCLLLVCRCTPTAPK----KTVTSRTG 559
                                                                                                                                                                                                                                                             Kawakami T., Sherman L., Dahlberg J., Gazit A., Yaniv A.,
Tronick S.R., Aaronson S.A.;
"Nucleotide sequence analysis of equine infectious anemia virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Coat protein, Glycoprotein; Polyprotein; Transmembrane; Signal.
SIGNAL 1 22 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (NOV-1987) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Equine infectious anemia virus (clone 1369) (EIAV) Viruses; Retroid viruses; Retroviridae; Lentivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ENV POLYPROTEIN.
COAT PROTEIN GP90.
COAT PROTEIN GP45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=87236196; PubMed=3035786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; C27842; VCLJEW.
InterPro: IPR001027; EIAV_GP45.
InterPro: IPR0011361; EIAV_GP90.
Pfam: PF010445; EIAV_GP45; I.
Pfam; PF00971; EIAV_GP95; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUL-1989 (Rel. 11, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M16575; AAB59863.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Virology 158:300-312(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GP90; COAT PROTEIN GP45]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REVISIONS TO N-TERMINUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              93
462
636
807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11670;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23
23
445
75
446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      614
787
816
40
112
1141
148
186
214
233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           368
399
406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         proviral DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ENV_EIAV9
P11306;
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          308 N--WQHRILLVRAREGM--------WIPVSTDRPWEASPSIHILTEIL 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53 CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122 CWGSFPGC--RPFQNY----FSYETNRSMHMDNNTATLLE-AYHREITFIYK 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    113 GK ------TCPK-EIPKGSKNTEVLV-------WEECVAN---SVVILQN 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   167 SSCTDSDHCQEYQCKKVNLNSSDSSNSVRVEDVTNTAEYWGFKWLECNQTENFKTILVPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           200 YLWEWEEKGIS-TPR----PKIISPVSGPEHPELWRLTVASHHIRIWSGNQTLETRYRK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-FYTIDLNSILTVPLQSCLKPPYMLVV---GNIVIKPASQTITC--ENCRLFTCIDSTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    146 NEFGTIIDX---APRGQFYHNCS---GQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                          Polyprotein; Transmembrane; Signal.
POTENTIAL.
ENV POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79; Mismatches 244; Indels 148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL
  us is not an important determinant of tropism in vitro."; Virol. 66:4085-4097(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.6%; Score 107; DB 1; Length 859; 20.7%; Pred. No. 0.62;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC. . .) (P. 23E020E80DF334FA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                    COAT PROTEIN GP90.
COAT PROTEIN GP45.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
POTENTIAL
                                                                                                                                                                                                                                        PIR; C41991; VCLJ22.
InterPro; IPR001027; EIAV GP45.
InterPro; IPR001361; EIAV_GP90.
Pfam; PF01045; EIAV_GP45; I.
Pfam; PF00971; EIAV_GP90; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              752 N
97140 MW;
                                                                                                                                                                                                                       EMBL; M87581; AAA43005.1; -.
                                                                                                                                                                                                                                                                                                                                          Coat protein; Glycoprotein; SIGNAL 1 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 20.7%
Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                    444
859
93
472
636
8907
835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 340
3368
3399
4406
411
483
483
550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥,
                                                                                                                                                                                                                                                                                                                                                                             233
233
233
445
617
617
787
780
1112
1141
1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 340
3368
3399
406
411
483
550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            254
      SO THE FERRE SERVICE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     а
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                       35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]
SEQUENCE FROM N.A.
MEDLINE-92292230; PubMed-1318398;
MEDLINE-92292230; PubMed-1318398;
Perry S.T., Flaherty M.T., Kelley M.J., Clabough D.L., Tronick S.R.,
Coggins L., Whetter L., Lengel C.R., Fuller F.;
"The surface envelope protein gene region of equine infectious anemia
                                                                                                                                                                                                                                                                                   306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           909
                                                                                                                                                                                                                                                          53 CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                                                      NEFGTIIDX---APRGQFYHNCS---GQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSF 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LKGVLNRSKRF-IFTLIAVIMGLIAVTATAAVAGVALH--SSVQSVNFVNYWQKNSTRLW 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NSQSSIDOK---LASQINDLRQTVIWMGDRLDLEHHFQL-QCDWNTSDFCITPQIYNESE 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) (CONTAINS: COAT PROTEIN
                                                                                                                                                                                                                                                                                                                                        113 GK----TCPK-EIPKGSKNTEVLV-------WEECVAN---SVVILQN
                                                                                                                                                                                                                                                                                                                                                                             SSCTDSDHCQEYQCKKVNLNSSDSSNSVRVEDVTNTAEYWGFKWLECNQTENFKTILVPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KP-FYTIDLNSILTVPLQSCLKPPYMLVV --- GNIVIKPASQTITC -- ENCRLFTCIDST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FGVIGQAHLELPRPNKRIRNQSFNQYNCSINNKTELETWKLVKT----SGVTPLPISSEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                200 YLWEWEEKGIS-TPR----PKIISPVSGPEHPELWRLTVASHHIRIWSGNQTLETR-YR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 330 VPIFYTCTFTNI----TSCNNEPIISVIMYETNOV-----OYLLCNNNNSNNYNCVVOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HPWNMSWGHLNESTQWDDWVSKMEDLNQEILTTLHCARNNLAQSMITFNTPDSIAQFGKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NGMDLIERQIKILYAMILQTHADVQLLKERQQVEETFNLIGCIERTHVFCHT-----G
                                                                                                                                                                                                                                                                                                                                                                                                                                             HHWDMVRRHL - - QGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNP -
                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          515 ----VTWIKTIRSTMIINLILIVVCLFCLLLVCRCTPTAPK-----KTVTSRTG 559
                    (POTENTIAL)
                                                                                                                                                                                                                 82; Mismatches 238; Indels 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LWSHIGNWIPGLGAS-IIKYIVWFLLIYLLL-----TSSPKILRALWKVTSGAG
                                                                                                                                                                            Query Match 3.6%; Score 107; DB 1; Length 859; Best Local Similarity 21.0%; Pred. No. 0.62; Matches 125; Conservative 82; Mismatches 238; Indels 1
N-LINKED (GLCNAC...) (PO

4B4ED8518CD4F364 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Equine infectious anemia virus (clone CL22) (EIAV). Viruses; Retroid viruses; Retroviridae; Lentivirus.NCBI_TaxID=31675;
  N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                     ΜW.;
                                                                                                                     97113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GP90; COAT PROTEIN GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
  4111
483
490
550
557
752
859 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ENV_EIAVC
P32541;
CARBOHYD
CARBOHYD
CARBOHYD
                                                                            CARBOHYD
CARBOHYD
                                                           CARBOHYD
                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345
                                                                                                                                                                                                                                                                                                                                                                                                                    146
                                                                                                                                                                                                                                                                                                                                                                                                                                                         227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ENV_EIAVC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
  FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                        õ
```

33;

Gaps

33;

```
CARBOHYD
                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    Query Match
                           CARBOHYD
                                                   CARBOHYD
                                                                                                       CARBOHYD
                                                                                                                                                          CARBOHYD
                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENV_SIVM1
AC BNO_S
AC BNO_S
DT 01-NC
DT 01-NC
DT 15-JT
DE ENVEL
DE ENVEL
CO SIMI.
OC VITU
OX VERU
CO (11)
CO (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       434
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Blooinformatics and the EMBL outstation the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                458
  | : | | : | : | CVIGQAHLELPRPUKRIRNQSFNQYNCSINNKTELETWKLVKT----SGVTPLPISSEAN 436
                                                                                                                             484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-APR-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) (CONTAINS: COAT PROTEIN
                                                                                                                             TGLIRHKRDFGISAIVAAIVAATAIAASATMSYVALTEVNKIMEVONHTFEVENST--LN
                                                                                                                                                                                   SQSSIDQK---LASQINDLRQTVIWMGDRLDLEHHFQL-QCDWNTSDFCITPQIYNESEH
                                                                                                                                                                                                                                    GMDLIERQIKILYAMILQTHADVQLLKERQQVEETFNLIGCIERTHVFCHT-----GH
                                                                                                                                                                                                                                                                                       HWDMVRRHL--QGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNP--
                                                                                                                                                                                                                                                                                                                                      PWNMSWGHLNESTQWDDWVSKMEDLNQEILTTLHGARNNLAQSMITFNTPDSIAQFGKDL
                                                                            KGVLNRSKRF-IFTLIAVIMGLIAVTATAAVAGVALH--SSVQSVNFVNYWQKNSTRLWN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MCGLIRE T.C., Lacy P.A., O'Rourke K.; "CDNA sequence of the env gene of a pathogenic equine infectious anemia lentivirus variant.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                            ----VTWIKTIRSTMIINLILIVVCLFCLLLVCRCTPTAPK-----KTVTSRTG 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                             608 WSHIGNWIPGLGAS-IIKYIVMFLLIYLLL-----TSSPKILRALWKVTSGAG 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Transmembrane; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Equine infectious anemia virus (strain WSU5) (EIAV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; Retroid viruses; Retroviridae; Lentivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COAT PROTEIN GP90.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ENV POLYPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             859 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nucleic Acids Res. 18:196-196(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=90174929; PubMed=2155398;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; X16988; CAA34856.1; -.
PIR; S07589; VCLJWS.
InterPro; IPR001027; EIAV_GP45.
InterPro; IPR001361; EIAV_GP90.
Pfam; PF01045; EIAV_GP45; 1.
Pfam; PF00971; EIAV_GP90; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GP90; COAT PROTEIN GP45]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         444
859
93
462
636
835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141
148
186
214
233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11671;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1
233
233
445
75
75
614
614
816
112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coat protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ENV_EIAVW
P16082;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
ENV_EIAVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                        381
                                                                            346
                                                                                                                                                                                403
                                                                                                                                                                                                                                    495
                                                                                                                                                                                                                                                                                       459
                                                                                                                                                                                                                                                                                                                                                                                            515
                                                                                                                                                                             ò
                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
-LNGMDLIERQIKILYAMILQTHADVQLLKERQQVEETFNLIGCIERTHVFCHT----- 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        53 CLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                                                                                                                                                                             -----WEECVAN---SVVILON 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        167 SSCTDSDHCQEYQCKKVNLNSSDSSNSVRVEDVTNTAEYWGFKWLECNQTENFKT1LVPE 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NEFGTIIDX---APRGQFYHNCS---GQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSF 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NEMVNINDTDTWIPKG-----CNETWARVKRCP-----IDILYGIHPIRLCVQPPF 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YLWEWEEKGIS-TPRPKIISPVSGPEHPELWRLTVASHHIRIWSGNQTLETRYRKPFYTI 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DLNSILTVP------LQSCLKPPYMLVV---GNIVIKPASQTITC--ENCRLFTCID 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DYTGIYQVPIFYTCTFTNITSCNNEPIISVIMYETNOV----QYLLCNNNNSNNYNCVV 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STFN--WQHRILLVRAREGM-----------WIPVSTDRPWEASPSIHILT 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QSFGVIGQAHLELPRPNKRIRNQSFNQYNCSINNKTELETWKLVKT----SGITPLPISS 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EILKGVLNRSKRF-IFTLIAVIMGLIAVTATAAVAGVALH--SSVQSVNFVNYWQKNSTR 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LWNSQSSIDOK---LASQINDLRQTVIWMGDRLDLEHHFQL-QCDWNTSDFCITPQIYNE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEHHWDMVRRHL--QGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLN 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P-----VIWIKTIRSTMIINLILIVVCLFCLLLVCRCTPTAPK-----KTVTSRTG 559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              605 KDLWSHIGNWIPGLGAS-IIKYIVMFLLIYLLL-----TSSPKILRALWKVTSGAG 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -GHPWNMSWGHLNESTQWDDWVSKMEDLNQEILTTLHGARNNLAQSMITFNTPDSIAQFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENV_SIVM1 STANDARD; PRT; 882 AA.
P05885; 085725; 085726;
01-NOV-1988 (Rel. 09, Careated)
01-NOV-1988 (Rel. 09, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
ENVELOPE POLYPROTEIN GP160 PRECIRSOR (CONTAINS: EXTERIOR MEMBRANE GLYCOPROTEIN (GP120); TRANSMEMBRANE GLYCOPROTEIN (GP41)].
                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                      (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                             Indels 154;
(POTENTIAL)
                                                                                                                                                                                 . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                     3.5%; Score 105; DB 1; Length 859; 20.6%; Pred. No. 0.91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Simian immunodeficiency virus (Mm142-83 isolate) (SIV-MAC)
Viruses; Retroid viruses; Retroviridae; Lentivirus.
                                                                                                                                                                                                                            6A7D8BC57050348D CRC64;
                                                                                                                                                                                                                                                                                                                             242;
                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                      (GLCNAC.
                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                  GK-----TCPK-EIPKGSKNTEVLV----
N-LINKED
                                                                                                                                                                                                                                                                                                                           78;
                                                                                                                                                                                                                            WW.
                                                                                                                                                                                                                          97106
                                                                                                                                                                                                                                                                                                                             Conservative
313
340
3340
3368
3399
4401
4411
4483
4490
752
                                                                                                                                                                               557
752
859 AA;
                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=11733;
[1]
313
340
368
399
406
411
483
550
                                                                                                                                                                                                                                                                                                                         Matches 123;
```

415

---KMNWFLNWV-E

491

415

695

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                              416 N----DLRQ----TVIWMGDRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHL 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWSGNQTLETRYRKPFYTIDLNSILT
                                                                                                                                                  26 VPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFNW----QHRILL-VRARE
                                                                                                                                                                                             ------VGKNVYLPPREGDLTCNS--TVTSLIANINWTDGNQTSITMSAEVAE
                                                                                                                                                                                                                                       321 GMWIPVSTDRPWEASPSIHILTEILK---GVLNRSKRFIFTL-----IAVIMGLIAV
                                                                                                                                                                                                                                                                                492 LYRLELGDYKLVEITPIGLAPTNVKRYTTGGTSRNKRGVFVLGFLGFLGFLATAGSAMGAASL
                                                                                                                                                                                                                                                                                                                           370 TATA----AVAGVALHSSVQSVNFVNYWQKNSTRL--WNSQ-----SSIDQKL--ASQI
                                                                                                                                                                                                                                                                                                                                                   NAWGCAFROVCHTTVPW------PNASLTPDWN------NETWOEWE---RKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QGREDNLT -- LDISKLKEQ -- IFEASKAHLNLVPGTEAIAGVADGLANLNPVT-WIKTIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-S288C / AB972;
MUTARAII Y. Naitou M., Hagiwara H., Shibata T., Ozawa M.,
Murakani Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
Yamazaki M., Tashiro H., Eki T.;
Analysis of the nucleotide sequence of chromosome VI from
Saccharomyces cerevisiae.";
Nat. Genet. 10:261-268(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98; Indels 162;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
HYPOTHETICAL 90.9 KDA PROTEIN IN GCN2O-CWK1 INTERGENIC REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3.5%; Score 104; DB 1; Length 787; 21.7%; Pred. No. 0.98;
                                                                                                       416 DRSLTTQKPK-----ERHKRNYVPCHIR-----QIINTWHK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90896 MW; 1919A8A1F58B5340 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           787 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46; Mismatches
                    380 NLTAPRGGDPEVTFMWTNCRGEFLYC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 648 DFLEANITALLEEAQIQQEKNMYELQK--LN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         523 --STMIINLILIVVCLFCLLLVCR 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  696 YGIYIIVGVILLRIVIYIVQMLAR 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D50617; BAA09252.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 3.5'
Best Local Similarity 21.7'
Matches 85; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          al protein.
787 AA; 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SGD; S0001909; I
Hypothetical pro
SEQUENCE 787 A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YFI3_YEAST
P43596;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YFR013W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                         610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YFI3_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   468
                                                                                                                                                                                          g
                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                               δy
                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---QAWCRFGGNWKEAIKEVKQTIVKHPRYTGTNNTD-------KI 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             43 ISIGYHYPPICLGRAPG--CLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQD 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  101 FSYQRSLKFRPKGKTCPKEIPK-----GSKNTEVLVWEECVANSVVILQNNEFGTI 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        152 IDXAPRGQ-----FYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLWEWE 205
                                    Chakrabarti L., Guyader M., Alizon M., Daniel M.D., Desrosiers R.C., Tiollais P., Sonigo P.; "Sequence of Simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses."; Nature 328:543-547(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISLNKHYNLTMKCRRPGNKTVLP------VTIMSALVFHS-OPVNERPK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.5%; Score 105; DB 1; Length 882;
20.0%; Pred. No. 0.94;
.ive 82; Mismatches 161; Indels 208;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLYCOPROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SRANE GLYCOPROTEIN
TERMINATION CODON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      216DD5C06C805DA7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IN-FRAME I
N-LINKED (N-LINKED (N-LIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                   EMBL; Y00277; -; NOT_ANNOTATED_CDS.
                    MEDLINE-87287230; PubMed=3649576;
                                                                                                                                                                                                                                                                                                                                                                                     PIR; H28887; VCLJG3.
HIV; M16403; ENV$MM142.
InterPro; IPR000328; Env_GP41.
InterPro; IPR000777; GP120.
Pfam; PF00516; GP120; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          101196 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 113; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        882 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

24;

Gaps

7PTVS 76 : :: [PLLT 155	110 /GLDL 214	5QFYH 162 FTE 268	sw 204 swroo 318	(PFYT 257 	FNWQHRILL 315 :: MDRVVL 391	
APGCLMPAVQNWLVE' : : VEPALIPSNWTSV	SYQRSLKFR :: : KFSSFFHSDLQNLSFQDFE	NNEFGTIIDXAPR ::: KMNLLFLSLLDLT	LW	IRIWSGNQTLETRYR : : IYTWNANEPLPLE-S	NCRLFTCIDSTFNWQ	
17 NDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNWLVEVPTVS 11	LMAFIN	111PKGKTCPKEIPKGSKNTEVLVWEECVA-NSVVILQNNEFGTIIDXAPRGGFYH 	NCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLW	205EEKGISTPRPKIISPVSGPEHPELMRLTVASHHIRIMSGNQTLETRYRKPFYT ::	IDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQHRILL :	LTEILK 346 HDEIYK 412
DDRCPAKPEEEGMMI KPAHE	PNSRFTYHMVSGMSLRPRVNYLQDF	PKEIPKGSKNTEVLV :: : : GIVKGPEDTSLLL	SAQVSPAVDSDLT SAKKKGPLTTW	RPKIISPVSG : : : KTALFEQDEQTPVVD	258 IDLNSILTVPLQSCLKPPYMLVV 	316 VRAREGMWIPVSTDRPWEASPSIHILTEILK 346 :
NDSVWVPGPTI 	PNSRFTYHMV: : :: SDFKNQYSVI		NCSGQTQSCP: : NFDGK	EEKGISTP :: LPSDQDISKP	IDLNSILT : REMDKNGIL-	VRAREGMWIP: LA
17	77 156	111	163	319	258	316
Oy Db	Qy	Qy Dp	Qy Db	Qy	Qy	Oy Dp

Search completed: April 9, 2002, 17:07:04 Job time: 568 sec

homo sapien homo sapien

```
095283 homo saplen
091400 homo saplen
071037 human endog
071072 human endog
071072 human endog
081394 mouse mamma
091210 endogenous m
085646 mouse mamma
091210 exogenous m
095145 exogenous m
095145 exogenous m
095145 mouse mamma
091177 ovis aries
09117 ovis aries
09117 ovis aries
09117 ovis aries
09117 ovis aries
0917 sheep pulmo
0917 ono sapien
015005 homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=9945825; PubMed=10516026; MEDLINE=9945825; PubMed=10516026; Toenjes R.R., Czauderna F., Kurth R.; Toenjes R.R., Czauderna F., Kurth R.; "Genome wide screening, clooning, chromosomal assignment and expression of full-length human endogenous retrovirus type K (HERV-K)."; J. VILO. 73:9187-9187-9195(1999). EMBL: Y18899; CABS5604.1; EMBL: Y18899; CABS5604.1; SEQUENCE 560 AA; 63671 MW; 8A4565663901BC3A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 2898.5; DB 12; Length 560;
Pred. No. 6.3e-261;
1; Mismatches 3; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 560 AA
                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=45617;
                           09HDC0
071037
0813194
081210
091217
091217
091217
091217
091217
091218
091218
091218
091218
091218
091218
091218
091218
091218
091218
091218
091716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       090C06 PRELIMINARY;
090C06;
01-MAY-2000 (TrEMBLrel. 13, C1
01-MAY-2000 (TrEMBLrel. 13, L6
01-MAR-2001 (TrEMBLrel. 16, L8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 99.1%;
Matches 542; Conservative 1
                                  382
368
368
363.5
363.5
362.5
360.5
341
340.5
222.5
222.5
223.5
183
777
751.5
621.8
491
472
395
393
393
389
388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
Q9QC06
 op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     O42043 unclassifie
O42043 unclassifie
Q9unw3 homo sapien
Q9unw2 homo sapien
O95284 homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9qc06 human endog
Q14773 homo sapien
Q9ubu4 homo sapien
Q63384 human endog
Q9yna5 human endog
Q9yna7 human endog
Q9ukh3 homo sapien
Q9ukh3 homo sapien
Q9uk77 homo sapien
Q9hx77 homo sapien
Q9ux77 homo sapien
Q9ux77 homo sapien
                                                                                                                  (without alignments)
1158.368 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                     1 MVTPVTWMDNPIEVYVNDSV......VCRCTPTAPKKTVTSRTGHE
                                                                                                 April 9, 2002, 17:06:31; Search time 70.84 Seconds
               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                             473505 segs, 146272329 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                        OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     090C06
014873
090B04
069384
087NA9
097NA7
097NA7
099B8
069386
069386
099DB8
069386
099DB8
                                                                                                                                                                                                                                 Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09UNW2
095284
095281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                         US-09-490-700-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                       SPTREMBL 17:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           153
153
153
153
153
                                                                                                                                                                                                                  BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     996.3
9911.6
9911.5
9911.5
9911.2
9911.2
9911.2
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
9911.3
991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9:
10:
11:
12:
13:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                        Database
                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3
6
7
7
7
8
8
111
111
113
114
118
118
118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
```

1;

```
834 LMPAVQNWLVEVPTVSPISRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEI 893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Meese E.;
                                                                                                                                                                                                                                                                                                                                       Q9UBU4;
                                                                                                                                                                                                                                                                                                                              Q9UBU4
                                                                                                                                                                                                                                                                                                            m
                                                                                                                         301
                                                                                                                                         1074
                                                                                                                                                                            1134
                                                                                                                                                                                                                                                                  540
                                                                                                                                                                                                                                                                                   1314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "An
                                                                     954
                                                                                      241
                                                                                                                                                           361
                 121
                                  894
                                                   181
                                                                                                                                                                                                                                                                                                                                                                                   EN
                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                      090BU4
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                  Dp
                ŏ
                                 QQ
                                                  ò
                                                                  9
                                                                                      ŏ
                                                                                                       a
                                                                                                                       ò
                                                                                                                                       g
                                                                                                                                                           ò
                                                                                                                                                                          g
                                                                                                                                                                                             ò
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                 ò
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1;
       Ç
                                           360
                                                    420
                                                                                               420
                                                                                                                479
                                                                                                                                  480
                                                                                                                                                           LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                   related
                                                                                      AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ
                                                                                                                TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS
                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                           TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI
                                                                                                                                                   KLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            91.6%; Score 2758.5; DB 4; Length 1361; 94.1%; Pred. No. 2.2e-247; ive 10; Mismatches 21; Indels 1;
                                                                                                                                                                                                                                                                                                                                                              MEDLINE-87036922; PubMed-3021993;
Ono M., Yasunaga T., Miyata T., Ushikubo H.;
"Nucleotide sequence of human endogenous retrovirus genome
the mouse mammary timor virus genome.";
                                                                                                                                                                                                                                                                                                                                                                                         the mouse mammary tumor virus genome.";
J. Virol. 60:589-598(1986).
-!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE TRANSCRIPTASE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CEB91B3F407B9498 CRC64;
                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                 PRT; 1361 AA
                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL, M14123; AAA88033.1; -.
InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_zn.
InterPro; IPR002156; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTSe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RNA-directed DNA polymerase.
SEQUENCE 1361 AA; 153797 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00552; integrase; 1.
Pfam; PF02022; Integrase_zn; 1.
                                                                                                                                                                                                                                                                 01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-UNN-2001 (TrEMBLrel. 17, POL/ENV ORF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF00075; rnaseH; 1
                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF00665; rve;
PF00078; rvt;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 515; Conserv
                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                    540 LLVCRCT 546
                                                                                                                                                                                               TISSUE-LIVER
                                                                                                                                                                                                                                                         014273;
                                                                                                                                                                                                                                                014273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                      541
                         241
                                           301
                                                           301
                                                                             361
                                                                                               361
                                                                                                                421
                                                                                                                                 421
                                                                                                                                                   480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                              qq
                                                      g
                                                                                              g
                                                                                                                            g
                                                                                                                                                                                   δ
                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                       g
                                         ò
                                                                            ò
                                                                                                               ò
                                                                                                                                                  ó
                                                                                                                                                                   Q
```

```
t
C
                                                                                                                                                                                                                                                                         TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI 360
                                                                                                                                                                                                                                                                                                                                                                 AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ 420
                                                                                                                                                                                WSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF 300
                   almost-intact human endogenous retrovirus K on human chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Lenz J.; "Many human endogenous retrovirus K (HERV-K) proviruses are unique
                                                                                                                                                                                                                                                                                                   PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV
                                                                                        DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Further characterization of the almost intact human endogenous retrovirus K on human chromosome 7."; Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases. EMBL; ARF04614; AAD51798 1; -.. EMBL; AF074086; AAF88168.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=99400989; PubMed=10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mayer J., Sauter M., Racz A., Scherer D., Mueller-Lantzsch N.
Meese E.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
Reus K., Mayer J., Sauter M., Scherer D., Mueller-Lantzsch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        699 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE=99178257; Pubmed=10080172;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genet. 21:257-258(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAR-2001 (TrEMBLrel. 16, ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    humans.";
Curr. Biol. 9:861-868(1999),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LLVCRCT 1320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LLVCRCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=9606;
```

ä

Gaps

ï

Indels

Length

231 180 291 240 351 300 411 360 471 420 531 479 591

171

9

```
LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                  SEQUENCE FROM N.A. Trochjes R.R., Cauderna F., Kurth R.; Tronjes R. R., Cauderna F., Kurth R.; Full-length human endogenous retrovirus type K (HERV-K) elements encoding Gag, Pol and Env proteins are localised on chromosomes 7
                                                       421 TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS
                                                                                                                                                                                                                                                                                                                                                                                       KLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCL
                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                       PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV
                                                                                                                                                                                     DSDLTESLDKHKKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI
                                                                                                                                                                                                192 DSDLTESLDKHKHKKKLQSFYPWEWGEKGISTPRPKIVSPVSGPEHPELWRLTVASHHIRI
                                                                                                                                                                                                                                                                                TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI
                                                                                                                                                                                                                                                                                                                            361 AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 12; Length 694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases EMBL; Y17832; CAA76880.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    80.1; -.
78690 MW; 82F91825669CF25B CRC64;
  DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               / Match 91.2%; Score 2746.5; DB 12; Local Similarity 93.8%; Pred. No. 1.2e-246; hes 513; Conservative 10; Mismatches 23;
91.5%; Score 2754.5; DB 1 94.0%; Pred. No. 2.1e-247;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  retrovirus K.
viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                     10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (TrEMBLrel. 10, (TrEMBLrel. 10, 1) (TrEMBLrel. 10, 1)
                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; Y17832; CAA761
SEQUENCE 694 AA;
             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human endogenous
Viruses; Retroid
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-1999 (
01-MAY-1999 (
01-MAY-1999 (
ENV PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 540 LLVCRCT
                        514;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Si
Matches 513;
 Query Match
              Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9YNA9
Q9YNA9;
                                                                                                                                                                                                                                                                                                                                                                                                                       480 1
                                                                                                                                                                                                                                                                                                                                                                                                 532
                                                                                         61
                                                                                                                                       121
                       Matches
                                                                                                                172
                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                               301
              Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
Q9YNA9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                g
                                                                                                                                                            g
                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                       ò
                                                                                                                                                                                    ò
                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                              ŏ
                                                                               1;
                                                                                                   DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI 240
                                                                                                                                                                                                                                                                   351
                                                                                                                                                                                                                                                                                          WSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF 300
                                                                                                                                                                                                                                                                                                     TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI 360
                                                                                                                                                                                                                                                                                                                                                   AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ 420
                                                                                                                                                                                                                                                                                                                                                                                                 TVIWMGDRL-DLEHHFOLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                       591
                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-95074858; PubMed-7983704;
Loewer R., Toenjes R.R., Korbmacher C., Kurth R., Loewer J.;
Loewer R., Toenjes R.R., Korbmacher C., Kurth R., Loewer J.;
Identification of a Rev-related protein by analysis of spliced
transcripts of the human endogenous retroviruses HTDV/HERV-K.";
J. Virol. 69:141-149(1995).
EMBL; X82272; CAA57723.1;
SEQUENCE 699 AA: 79217 MW; 3844DA6D80D76DFC CRC64;
                                                                               1;
                                                        4; Length 699;
                                                                              Indels
                       3844DA6D80D76DFC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                              22;
                                                       ; Score 2754.5; DB 4;
; Pred. No. 2.1e-247;
10; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human endogenous retrovirus.
Viruses; Retroid viruses; Retroviridae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                       79217 MW;
                                                       91.5%;
94.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel. 01, 0
(TrEMBLrel. 01, 1)
(TrEMBLrel. 08, 1
EMBL; AF074086; AAD21098.1;
                                                      Query Match 91.5%
Best Local Similarity 94.0%
Matches 514; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
            Envelope protein.
SEQUENCE 699 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11827;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ||||||||||
LLVCRCT 658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1996 (
01-NOV-1996 (
01-NOV-1998 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LLVCRCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          069384
069384;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4
                                                                                                                           112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            540
                                                                                                      -
                                                                                                                                                                         172
                                                                                                                                                                                               121
                                                                                                                                                                                                                      232
                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                         472
                                                                                                                                                                                                                                                                                                                                                                                                                                                       532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  652
                                                                                                                                                  61
                                                                                                                                                                                                                                                                  292
                                                                                                                                                                                                                                                                                                              352
                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                              412
                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ENV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
Q69384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SW
                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                           유
                                                                                                                                                                                               ò
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                      à
                                                                                                                                                  ð
```

651

ä

Gaps

;;

Indels

9

and

360 466 420 646

107 61 167 121 227 181 287 241 347 301 407 361 467 421 527 480

> QQ à 요 ò qq ò

```
[1]
SEQUENCE FROM N.A.
Toenjes R.R. (Zauderna F., Kurth R.;
Toenjes R.R. (Zauderna F., Kurth R.;
"Full-length human endogenous retrovirus type K (HERV-K) elements
"Full-length human endogenous retrovirus type K (HERV-K) elements
encoding Gag, Pol and Env proteins are localised on chromosomes 7.
                                                                              IMPAVQNWLVEVPTVSPICRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEI
                                                                                                                                                                                                      AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ
                                                                                                                                                                                                                                                        TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS
                                                                                                                                                                                                                                                                                                             MYTPYTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV
                                                                                                                                   WSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF
                                                                                                                                                                                      TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI
                                                                                                                                                                                                                                                                                                                                                  KLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 2733.5; DB 12; Length
Pred. No. 1.9e-245;
0; Mismatches 25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases EMBL; Y17833; CAA76883.1; •. SEQUENCE 694 AA; 78692 MW; 498A1D745366D0EE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90.8%; Scilarity 93.4%; Pr
Conservative 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (TrEMBLrel. 10, (TrEMBLrel. 10, (TrEMBLrel. 13,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                      LLVCRCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                     LLVCRCT (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-1999 (
01-MAY-1999 (
01-MAY-2000 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Simi
Matches 511;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9YNA7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9YNA7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7
 167
                                                                                                                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                                                                                                      540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                107
                                                     227
                                                                                                        287
                                                                                                                                   241
                                                                                                                                                            347
                                                                                                                                                                                      301
                                                                                                                                                                                                                 407
                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                             421
                                                                                                                                                                                                                                                                                                                        527
                                                                                                                                                                                                                                                                                                                                                                                                                               647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         227
                          121
                                                                               181
                                                                                                                                                                                                                                                                    467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ENV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
Q9YNA7
                                                  g
                                                                                                     qq
                                                                                                                                                            g
                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                         ò
                                                                                                                                 ŏ
                                                                                                                                                                                      δ
                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ä
                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.

Toenjes R.R., Czauderna F., Kurth R.;
"Full-length human endogenous retrovirus type K (HERV-K) elements
encoding Gag, Pol and Env proteins are localised on chromosomes 7 and
                                       120
                                                                                                                     286
                                                                                                                                                 240
                                                                                                                                                                                                                             406
                                                                                                                                                                                                                                                                                                                                                              479
                                                                                                                                                                                                                                                                                                                                                                             LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                          PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                                                                   WSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS
                                                                                                                                                 DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI
                                                                                                                                                                                                                                                                                                                                                                                                                   KLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCL
                                       LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91.0%; Score 2741.5; DB 12; Length 694; 93.6%; Pred. No. 3.4e-246; ive 11; Mismatches 23; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases EMBL; Y17834; CAA76886.1; -. SEQUENCE 694 AA; 78635 MW; 23A8A5A188543317 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=45617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 10, C
(TrEMBLrel. 10, I
(TrEMBLrel. 10, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
les 512; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LLVCRCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LLVCRCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9YNA5
Q9YNA5;
Q1-MAY-1999 (
Q1-MAY-1999 (
ENV PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loc
Matches
```

647

g

ò

a

ò a g

à

RESULT Q9YNAS

Æ. 694

ij

Gaps

ij

107 61

g ò S

õ 셤 ò ద ò g ò g ò a ò g ò g

```
ដ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -i- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
-i- SIMILARITY: THE PROTEASE BELONGS TO THE PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EUKARYOTIC ASPARTYL PROTEASES FAMILY.
EMBL; API64611; AAD51793.1; -.
HSSP; P16088; 1DUT.
                                                    360
                                                                                                                    470
                                                                                                                                             420
                                                                                                                                                                                                                                                      539
300
                                                                                                                                                                                                  TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS 479
                                                                                                                                                                                                                                                                                 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Many human endogenous retrovirus K (HERV-K) proviruses are unique
                                                                                          TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI
                                                                                                         WSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF
                                                                                                                                              361 AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ
                                                                                                                                                                                                                                                      KLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE=99400989; Pubmed=10469592;
Barbulescu M., Turner G., Seaman M.I., Delnard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Curr. Biol. 9:861-868(1999).
-!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
TRANSCRIPTASE).
                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 2294 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR001969; Asp_protease.
Interpro; IPR001995; Asp_prot_retrov.
Interpro; IPR001428; duTpase.
Interpro; IPR003323; Gag_D10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Integrase_zn.
RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Integrase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF00552; integrase; 1.
PF02022; Integrase_Zn; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000477; RVTse.
InterPro; IPR001878; Znf_CCHC.
Pfam; PF00692; dUTPase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PP00692; durpase; 1.
PP01585; G-patch; 1.
PP02337; Gag_p10; 1.
PP00607; gag_p24; 1.
PP00552; integrase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gag_p24.
G_patch.
                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17, GAG-PRO-POL-ENV PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000721; GallerPro; IPR000467; GallerPro; IPR001037; II InterPro; IPR003308; II InterPro; IPR002156; R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR001584;
                                                                                                                                                                                                                                                                                                          LLVCRCT 546
                                                                                                                                                                                                                                                                                                                                    657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                       LLVCRCT (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lenz J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    humans.
                                                                                                                                                                                                                                                                                                                                                                                                                   09UKH9;
                                                                                                                                                                                                                                                                                                                                                                                                     090кн9
                                                                                                                                                                                                                                                                                                                                                                            σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                          301
             291
                                      241
                                                                                                                                                                       471
                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                      480
                                                                                                                                                                                                                                                                                                          540
                                                                                                                                                                                                                                                                                                                                   651
                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                         о9икн9
                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                               g
                                                                                        à
                                                                                                                 g
                                                                                                                                              à
                                                                                                                                                                      q
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5
                                     346
                                                            420
                                                                                                                                                                                                                           479
                                                                                                                                                                                                                                                                                             ಭ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Lenz J.; "Many human endogenous retrovirus K (HERV-K) proviruses are unique
           TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI
                                                                                                                                                                                                                           TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDIS
                                                                                                                                                                                                                                                                              KLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCL
                                                                                                                                                                       AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalja; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDIINE-99400989; PubMed-10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd K.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.8%; Score 2705; DB 4; Length 698; 93.1%; Pred. No. 8.5e-243; 1.ve 11; Mismatches 25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Curr. Biol. 9:861-868(1999).
EMBL, AFI44615; AAD51800.1; -.
SEQUENCE 698 AA, 79015 MW, 1423B81945C87BAD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                               698 AA
                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 93.19
Matches 509; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                                                                                                   LLVCRCT 546
                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                             о9икнз;
                                                                                                                                                                                                                                                                                                                                                                                                                              о9икн3
                                                                                                                                                                                                                                                                                                                                                                                                    œ
                                                                                                                                                                                                                                                                                                                                   240
                                                                                                                                                                                                                                                                              480
                                                                                                                                                                                                                                                                                                        587
                                                                                                                                                                                                                                                                                                                                                           647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
            181
                                     287
                                                                                                                 301
                                                                                                                                            407
                                                                                                                                                                      361
                                                                                                                                                                                               467
                                                                                                                                                                                                                         421
                                                                                                                                                                                                                                                  527
                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
Q9UKH3
```

g

ò

ò g a à

ð

ø

ö

ò g ò g ò g ò a ò g ŏ qq ò

```
mammalian cells
                                                                                                                                         67.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40.3%;
96.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16,
 glycoprotein in insect and r
J. Virol. 71:2747-2756(1997)
                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                         Query Match
Best Local Similarity
Matches 376; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Simil
Matches 222; C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09нрв8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09HDB8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11
                                                                                                                                                                                                                   112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           472
                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                      232
                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                       292
                                                                                                                                                                                                                                                                                                                                                                                                241
                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
 qq
                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                          ò
                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                   ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                              2215 TVIMMGDRLMSLEHRFQLQCDWNTSDFCITPQIYNESEHHWDWYRRHLQGREDNLTLDIS 2274
                                                                                                                                                                                                                                                                                                                                                                                                240
                                                                                                                                                                                                                                                                                            LMPAVQNWLVEVPTVSPNSRFTYHWVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361 AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMYRRHLQGREDNLTLDIS 479
                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                241 WSGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF
                                                                                                                                                                                                                                           1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI
                                                                                                                                                                                                                   1;
                                                                                                                                                                                          Length 2294;
R Pfam; PF00075; rve; 1.

R Pfam; PF00077; rve; 1.

R Pfam; PF00078; rvt; 1.

R Pfam; PF000945; duTPase; 1.

SMART; SM00343; doTPatch; 1.

R RARRT; SM00343; ZnF_C2HC; 2.

R R RPGSTTE; PS00141; ASP_ENOTERES; 1.

R PROSTTE; PS01145; ASP_PROT_RETROY; 1.

R ASPARTY] protease, Hydrolase, RNA-directed DNA polymerase; 1.

R Aspartyl protease, Hydrolase, RNA-directed DNA polymerase; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.

BEDLINE-97213939; PubMed-9060628;

Tonjes R.R., Limbach C., Lower R., Kurth R.;

"Expression of human endogenous retrovirus type K envelope
                                                                                                                                                                                                                   24; Indels
                                                                                                                                                   2294 AA; 257099 MW; 425C81FDDF8A80D0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                         81.5%; Score 2452.5; DB 4; 92.9%; Pred. No. 1.3e-218;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                514 AA
                                                                                                                                                                                                     Best Local Similarity 92.9%; Pred. No. 1.3e
Matches 457; Conservative 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human endogenous retrovirus K.
Viruses; Retroid viruses; Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1998 (TrEMBLrel. 08, POL/ENV GENE (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2275 KIKRTNFRSIKS 2286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       480 KLKEQIFEASKA 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=45617;
                                                                                                                                                     SEQUENCE
                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q69386
Q69386;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
069386
AC 069386
AC 069386
DT 01-NOV
DT 01-NOV
DT 01-NOV
DC NOVER
CN POL/EN
CN POL/EN
CN VITUSE
CN NORL_T
RN REDLIN
RX MEDLIN
RX REDLIN
RX "Expre
                                                                                                                                                                                                                                                                   1795
                                                                                                                                                                                                                                                                                                                                                                                                                       1975
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2155
                                                                                                                                                                                                                                                                                                                      1855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2095
                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
```

ò

```
LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                           PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WSGNOTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASQTITCENCRLF 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DSDLTESLDKHKKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Jinno Y., Sugimoto J.; "Human endogenous retrovirus HERV-K(II) and flanking sequences."; "Human endogenous retrovirus HERV-K(II) and flanking sequences."; Embil: AB047240; BAB11/60.1; -. SEQUENCE 245 AA; 27904 MW; 198F26D65ED56DDB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                 ;
0
                                                                                                                                                          514;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                       Length
                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
Toenjes R.R.;
Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
EMBL; X92887; CAA63481.1; -.
NON TER 514 514
SEQUENCE 514 AA; 58200 MW; 86410F35B82DF231 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361 AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNS 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                          DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 1213; DB 4;
Pred. No. 1.2e-104;
1; Mismatches 8;
                                                                                                                                                     Score 2028; DB 12;
Pred. No. 5.3e-180;
7; Mismatches 20;
```

ó

ö

,77XN90 **09NX77**

RESULT 12

09NX77

121 181 181

61

ð g ò g ð g

```
S
                                                                                                                                                                                                                                                                                                                                                                                           61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                            a CA repeat
                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                              "Many human endogenous retrovirus K (HERV-K) proviruses are unique
                                                                                                                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Buteleostoml;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Rao J.,
                                                                                                                                                                                                                                                            Length 277;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 153;
                                                                              MEDLINE-99400989; PubMed-10469592;
Barbulescu M., Turner G., Seaman M.I., Deinard A.S., Kidd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13 PKESKNTEVLVWEECVANSAVILQNNEFGTIIDWAPRGQFYHNCSG 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. MEDDMed=9778243; MEDLINE=98449692; PubMed=9778243; Lan M.S., Mason A., Coutant R., Chen Q.Y., Vargas A., Gomez R., Chalew S., Garry R., Maclaren N.K.; "HERV-KIOS and immune-mediated (type 1) diabetes.";
                                                                                                                                                                                                            E2FCF565CCFA85E5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       . 153 AA; 17305 MW; 2B36E1656DF0C01A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27.9%; Score 840; DB 4; I
100.0%; Pred. No. 3.2e-70;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                            Score 845; DB 4;
Pred. No. 2.5e-70;
                                                                                                                                                                                                                                                                                          2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                            277 AA; 31420 MW;
                                                                                                                                                                                                                                                             28.1%;
92.8%;
                                                                                                                                                          Curr. Biol. 9:861-868(1999).
EMBL; AF164612; AAD51795.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-NOV-1999 (TrEMBLrel. 12, ENVELOPE PROTEIN RIC-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                            Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cell 95:14-16(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                Envelope protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Envelope protein.
SEQUENCE 153 AA
                                NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Sim
Matches 153;
                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            095280;
                                                                                                                                                humans.
                                                                                                                Lenz J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            095280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
P
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRLWNSQSSIDQKLASQINDLRQTVIWMGDR-LDLEHHFQLQCDWNTSDFCITPQIYNE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAV 180
                                                                                               277 MLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASP 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKN 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPV 515
               LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                    LEGUENCE FROM N.A.
TISSUE-ILEAL MUCOSA;
TISSUE-ILEAL MUCOSA;
Tanigami A., Fujiu Y., Ozaki K.,
HIRAO M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
NEDO human CDNA sequencing project.",
Submitted (FEB-2000) to the BMEL/GenBank/DDBJ databases.
EMBL; AKO00400; BAA91140.1; --
SEQUENCE 271 AA; 30693 MW; ECIS90F3D599660E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 271;
                                                                                                                                           49; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 31.7%; Score 955.5; DB 4; Best Local Similarity 64.4%; Pred. No. 1.3e-80; Matches 174; Conservative 46; Mismatches 49;
                                                                                                                                                                                                                                                            271 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-0CT-2000 (TrEMBLrel. 15, Created) 01-0CT-2000 (TrEMBLrel. 15, Last seq 01-0CT-2000 (TrEMBLrel. 15, Last ann CDNA FLJ20393 FIS, CLONE KAIA4699.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo saptens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
```

ö

090КН7 090КН7;

RESULT 13 Q9UKH7

8

õ g

337 61 397

ö 셤 õ g ö g ò

```
;
SEQUENCE FROM N.A.

MEDLINE=97386455; PubMed=9244304;

Conrad B., Weissmahr R.M., Boni J., Arcari R., Schupbach J., Mach B.;

Conrad B., Weissmahr R.M., Boni J., Arcari R., Schupbach J., Mach B.;

"A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes.";

Cell 90:303-313(1997).

EMBL; AF012337; AAC58456.1; -.

EMBL; AF012336; AAC58456.1; -.

ENDL; AF012336; AAC58456.1; -.

SEQUENCE 153 AA; 17305 MW; 2B36E1656DF0C01A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 27.9%; Score 840; DB 12; Length 153; Best Local Similarity 100.0%; Pred. No. 3.2e-70; Matches 153; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           153 AA
                                                                                                                                                                                  Viruses; Retroid viruses; Retroviridae.
NCBI_TaxID=35276;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JAN-1998 (TrEMBLrel. 05, Created)
01-MN-1998 (TrEMBLrel. 05, Last sequ
01-NV-1998 (TrEMBLrel. 08, Last anno
ENVELOPE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unclassified Retroviridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 042043
                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
042043
AC 042043
AC 042043
DT 01-JAN
                                                                                      g
                                                                                                                                                                                            ð
```

61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120

121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153

121

ò

Search completed: April 9, 2002, 17:06:33 Job time: 582 sec

g ŏ QQ

õ

Human protein sequ Human protein sequ Peptide #496 encod Peptide #495 encod Peptide #474 encod

Jaagslekte

AAR31349 AAB95842

Human secreted pro Extended human sec

AAM14052 AAM26458 AAM01792 AAG03399 AAY35898 AAY1086 AAB15724

Human lectomedin-1 Human lectomedin-1 Peptide Seg ID No: Protein encoded by

Perfect score:

Sednence:

OM protein

Run on

Scoring table:

Searched:

Database

```
IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human endogenous retrovirus IDDMK1.2-22 env/fs-sag protein.
                                                                                                                                                                                                                                                                                                                                                                    AAX41091
AAB15729
AAB15729
AAB67114
AAB67126
AAB67126
AAB67126
AAB67129
AAM14987
AAB67120
AAB67120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW95693 standard; Protein; 561 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97EP-0401773
 (first entry)
 WPI; 1999-097928/09.
N-PSDB; AAX07515.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mach B;
(MACH/) MACH B F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EP893691-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     27-JAN-1999
Conrad B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW95693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW95693
 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide #1333 enco
Peptide #1365 enco
Peptide #1305 enco
Human endogenous r
                                                                                                                                 (without alignments)
574.362 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human endogenous r
Human endogenous r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human endogenous r
Human endogenous r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human endogenous r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                  1 MVTPVTWMDNPIEVYVNDSV......VCRCTPTAPKKTVTSRTGHE
                                                                                                              April 9, 2002, 16:56:43; Search time 72.35 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  **Controlled **Con
               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
                                                                                                                                                                                                                                                                                                                                      hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                     522463 seqs, 74073290 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                              - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW97746
AAM14899
AAM27328
AAM02623
AAW95692
AAW95692
AAW95694
AAW97747
AAW97745
AAW97745
                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                 us-09-490-700-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A_Geneseq_1101:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      561
312
312
312
312
153
153
181
181
181
688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  999.9
999.9
444.0
227.9
227.9
118.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3009
3009
1325.5
1325.5
1325.5
840
840
840
840
840
840
840
840
```

Sequence homologou Rat latrophilin pr Rat latrophilin. Calcium independen Mouse mammary tumo Human mammary tumo

Cat mammary tumour

Sequence encoded b EIAV env gene prod

ALIGNMENTS

AAB35756

Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells,

Human protein sequ

Score

Result Š Mouse mammary tumo

Protein derived fr

Rhesus mammary Human mammary

Human lectomedin-1 Human lectomedin-1 Human lectomedin-1 Human lectomedin-1 Rhesus mammary tum

AAB15730 AAY41085 AAX41085 AAX41087 AAB15725 AAB67118 AAB67115 AAB67115 AAB67127 AAB67125 AAB67125

tumo tumo tumo

Human mammary Human mammary Human mammary (first entry)

21-MAY-1999

AAW97746;

```
The section of many control of the c
specific binding agents, used for treating
                                                                                                                                                                                                                                              an insulin-dependent diabetes mellitus
                                        preventing autoimmune disease
                                                                                                                                         Claim 31; Fig 7E; 92pp; English.
proteins, antibodies and
                                                                                                                                                                                                                                              sequence is that of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           561 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
```

Score 3009; DB 20; Length 561; Pred. No. 1e-283; 0; Mismatches 1; Indels 0. 99.98; Matches 560; Conservative Similarity Query Match Best Local S

·; LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120 Gaps 9 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC ö 61 61 121 121 3/8 g ò qq

240 300 360 TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI 360 420 AVIMGLIAVTATAAVACVALHSSVQSVNFVNVWQKNSTRLWNSQSSIDQKLASQINDLRO DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI WSGNOTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVGNIVIKPASOTITCENCRLF 181 181 241 241 301 301 361 ò Dβ g G ŏ Q à

TVIWMGDRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISK LKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCLL 421 421 481 ò 원 ă

480

420

361

g

원

ò 셤

541

AAW97746 standard; Protein; 561 AA. AAW97746 ID AAW9

1

```
This is the envelope protein (Env) of a new human endogenous retrovirus (HERV), designated IDDMK1.2-22, that has been identified as the source of superantigen (SAg) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAg activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV having SAg activity comprises specifically detecting in a biological sample either: (a) the mRNA of an expressed HERV having SAg activity (especially IDDMK1.2-22 5'LTR, 3' long terminal crepeat, env or pol); (b) a protein or peptide expressed by the HERV (see AAM97745-48); (c) antibodies specifically associated with the CHERV. products of the invention can be used to identify substances capable of blocking transcription or translation of SAg-encoding nuclear acid sequences, useful in therapy and/or prevention of human proteonical acid sequences useful in therapy and/or prevention of human proteonical acid sequences useful in the source of the force of the human crises associated with the SAg. A nucleic acid encoding the protein of the source of the force of the f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            auroimmune disease associated with the SAq. A nucleic acid encoding human retroviral SAq can be used as a DNA vaccine. Expression of the endogenous SAq in IDDM suggests a general model according to which self SAq-driven and systemic activation of autoreactive T
                                                                                                                                                                                 HERV; IDDKK1.2-22; superantigen; SAg; antigen; IDDM;
insulin-dependent diabetes mellitus; autoimmune disease; diagnosis;
therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New isolated human endogenous retrovirus - used to develop products for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                              Human endogenous retrovirus IDDKK1.2-22 envelope protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cells leads to organ-specific autoimmune disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                42..45
/note= "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                             17..19
/note= "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 6; Fig 7E; 165pp; English.
                                                                                                                                                                                                                                                                                              Human endogenous retrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97EP-0112482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-143118/12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                    Key
Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9905527-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-FEB-1999
```

Gaps ö Length 561; Indels DB 20; 1; Score 3009; DB 20 Pred. No. 1e-283; 0; Mismatches 99.98; Conservative Similarity Best Local Sim Matches 560; Query Match

Ä 561

Sequence

ó

1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60

ò Q

```
Sequence
                                                                                                                                                                                                                                                                                                                                   AAM27328
                                                                                                                                                                                                                                                                                                                                                                            Probe;
                   Claim
                                                                                                                                                                                        337
                                                                                                                                                                                                      61
                                                                                                                                                                                                                                              456
                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                   397
                                                                                                                                                                                                                                                                                                               AAM27328
                                                                                                                                                                           g
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                             ò
                              Peptide #1333 encoded by probe for measuring cervical gene expression
                                                                       dsdltesldkhkkklqsfylweweekgistprpkiispvsgpehpelwrltvashhiri 240
                                                                                    300
                                                                                           TCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFTLI 360
                                                                                                                     AVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDLRQ 420
                                                                                                                                                 TVIWMGDRLDLEHHFOLOCDWNTSDFCITPOIYNESEHHWDMYRRHLOGREDNLTLDISK 480
                                                                                                                                                                                               LKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLILIVVCLFCLL 540
                                                                                                                                                                                                                                                                                                                             human; microarray; gene expression; cervical epithelial cell;
            WSGNQTLETRYRKPFYTIDLNSILTVPLOSCLKPPYMLVVGNIVIKPASOTITCENCRLF
                                                                                                                                                                            LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKE1
                                                          DSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRI
                                                                DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rank
                                                                                                                                                                                                                                                                         AA.
                                                                                                                                                                                                                                                                       AAM14899 standard; Protein; 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                         LVCRCTPTAPKKTVTSRTGHE 561
                                                                                                                                                                                                                                Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0207456.
2000US-0608408.
2000US-0632366.
2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                              2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                        30-JAN-2001; 2001WO-US00670
                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0236359
                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-488901/53
                                                                                                                                                                                                                                                                                                                                   cervical cancer
                                                                                                                                                                                                                                                                                                                                                             WO200157278-A2
                                                                                                                                                                                                                                                                                                                                                                                                           26-MAY-2000;
30-JUN-2000;
03-AUG-2000;
21-SEP-2000;
27-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                     04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                              04-OCT-2000;
                                                                                                                                                                                                                                                                                                  12-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                          09-AUG-2001
                                                                                                                                                                                                                                                                                    AAM14899;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Penn SG,
                                                                                                                                                                                                                                                                                                                             Probe;
                                                                                                                                                                                                                         541
                 61
                                           121
                                                                                    241
                                                                                                 241
                                                                                                              301
                                                                                                                                         361
                                                                                                                                                                    421
                                                                                                                                                                                              481
                                                                                                                                                                                                                                                                                                                                                 Homo
     61
                                                         181
                                                                       181
                                                                                                                           301
                                                                                                                                                      361
                                                                                                                                                                                 421
                                                                                                                                                                                                                                                                 AAM14899
                                                                                                                                                                                                                                                                              g
                                                                      g
                                                                                                g
                                                                                                                          g
                                                                                                                                                     g
                                                                                                                                                                                a
                                                                                                                                                                                             à
                                                                                                                                                                                                            셤
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                     g
                                          g
                                                                                   ò
                                                                                                              õ
                                                                                                                                       ò
                                                                                                                                                                   ò
                                                        ò
```

```
ï
                                                                                                           The present invention relates to human single exon nucleic acid probes (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably
                                                                                                                                                                                                                                                                         cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide #1365 encoded by probe for measuring placental gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKN 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              277 MLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPV
for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22; Length 312;
                     analyzing gene expression in human cervical epithelial cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1325.5; DB 22; Lengt
Pred. No. 2.5e-120;
6; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               microarray; human; placenta; antenatal diagnosis;
genome-derived single exon nucleic acid
                                                                 English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          516 TWIKTIRSTMIINLILIVVCLFCLLLVCRCT 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241 twoktigsttiinlillilloclfclllocrct 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ą
                                                               ID NO 19725; 487pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAM27328 standard; Protein; 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0180312.
2000US-0207456.
2000US-0608408.
2000US-023366.
2000US-0234687.
2000US-0234587.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  44.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2001WO-US00663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 44.0
Best Local Similarity 93.7
Matches 254; Conservative
                                                                                                                                                                                                                                                                                                                                                                                              312 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genetic disorder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200157272-A2
                                                                 SEO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     04-FEB-2000;
26-MAY-2000;
30-JUN-2000;
03-AUG-2000;
21-SEP-2000;
27-SEP-2000;
04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-JAN-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     17-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09-AUG-2001
                                                               27;
```

```
EP893691-A1
                                                                                                                                                                                                                                                                                           Matches 254;
                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAW95692;
                                                                                                                                                                                                                                                                         Query Match
                                          SG,
                                                                                                                                                                                                                                                                                   Local
                                                                            Novel
                                                                                                                                                                                                                                                                                                                                            337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241
                                           Penn
                                                                                                                                                                                                                                                                                                                                                                              397
                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                 456
                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW95692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                             δý
                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Probe; human; breast disease; breast cancer; development disorder; inflammatory disease; proliferative breast disease; non-carcinoma tumour.
                                                                                                    The present invention relates to single exon nucleic acid probes (SENP: see AAI31315-AAI57546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of
                                                                                                                                                                                                                                                                                                                                   515
                                                                                                                                                                                                                                                                          STRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNE 455
                                                                                                                                                                                                                              277 MLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASP 336
                                                                                                                                                                                                                                                                 SIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKN 396
                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptide #1305 encoded by probe for measuring breast gene expression
                                                                                                                                                                                                                                        SEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPV
                                                          genome-derived single exon nucleic acid probes useful for
                                                                                                                                                                                            DB 22; Length 312;
                                                                                                                                                                                                             10; Indels
                                                                                                                                                                                                    Pred. No. 2.5e-120;
6; Mismatches 10;
                                                                    analyzing gene expression in human placenta
                                                                                                                                                                                           44.0%; Score 1325.5;
93.7%; Pred. No. 2.5e
                                                                                  SEQ ID No 27597; 654pp; English.
                                                                                                                                                                                                                                                                                                                                                                    516 TWIKTIRSTMIINLILIVVCLFCLLLVCRCT 546
                                                                                                                                                                                                                                                                                                                                                                               Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                 AAM02623 standard; Protein; 312 AA
       (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0207456.
2000US-0608408.
2000US-0632366.
2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2001WO-US00661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                            Matches 254; Conservative
                                                                                                                                                 human genetic disorders.
                                         WPI; 2001-488897/53
                                                                                                                                                                                                    Similarity
                                                                                                                                                                 312 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200157270-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-FEB-2000; 26-MAY-2000; 30-JUN-2000; 03-AUG-2000; 21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-JAN-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-AUG-2001
                                                                                   Claim 27;
                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAM02623;
                                                                                                                                                                                             Query Match
                        SG,
                                                                                                                                                                                                      Local
                        Penn
                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                 337
                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                  397
                                                                                                                                                                                                                                                                                                                                   456
                                                                                                                                                                                                                                                                                                                                                                                                                         AAM02623
ò
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                  å
                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                   a
```

```
(see AA10010-AA110067). The present sequence is a peptide encoded by one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probe hybridises at high stringency to a nucleic acid expressed in the human breast. The probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aeticlogy. The diseases include breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               printed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              breast disease and non-carcinoma tumours.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                /el single exon nucleic acid probe used to measuring gene expression a human breast {\boldsymbol{\cdot}}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITPQIYNE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       277 MLVVGNIVIKPASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASP 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKN 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IDDM; insulin-dependent diabetes mellitus; endogenous retrovirus; SAg; superantigen; provirus; autoimmune disease; type 1 diabetes; diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                                                                                                                                                relates to novel single exon nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human endogenous retrovirus IDDMK1.2-22 env protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 1325.5; DB 22
Pred. No. 2.5e-120;
6; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         at ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                Claim 27; SEQ ID No 11363; 322pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          516 TWIKTIRSTMIINLILIVVCLFCLLLVCRCT 546
                                                                                                                                               DR;
                                                                                                                                               Rank
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAW95692 standard; Protein; 153
                                                                                     (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                               Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44.0%;
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JUN-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                present invention
                                                                                                                                               Hanzel DK,
                                                                                                                                                                                                      WPI; 2001-476286/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
```

ï

S

```
Human endogenous retrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
ò
                                                                                                                                                                                                                                                                                                                                                                              The sequence is that of an insulin-dependent diabetes mellitus associated human endogenous retrovirus (IDDMK1.2-22) env protein. The retrovirus has Superantique (SAg) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence (without SAg activity but still able to induce an immune response) are useful in vaccines to treat or prevent SAg-related autoimmune disease; nucleic acid sequences encoding (modified) SAg are important in pathogenesis of autoimmune disease, probably by activating to treat such disease, probably by activating autoractive T cells. The method is very specific (it can differentiate between expressed and non-expressed viral nucleic acids) and can be used even where the pathogen is an ubiquitous endogenous retrovirus. Blood cor plasma samples can be tested without extensive preparation and diagnosis can be made before clinical signs are apparent, allowing early intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating or preventing autoimmune disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HERV; IDDXK1.2-22; superantigen; SAg; antigen; IDDM;
insulin-dependent diabetes mellitus; autoimmune disease; diagnosis;
therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human endogenous retrovirus IDDKK1.2-22 envelope protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27.9%; Score 840; DB 20;
100.0%; Pred. No. 1.4e-73;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 pkgskntevlvweecvansvvilgnnefgtiid 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAW97745 standard; Protein; 153
                                                                                                                                                                                                                                                                                                                                                    Claim 31; Fig 7D; 92pp; English.
                                               97EP-0401773
                                                                                  97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                         WPI; 1999-097928/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                        Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153 AA;
                                                                                                                     (MACH/) MACH B F.
                                                                                                                                                                                                          N-PSDB; AAX07514.
                                               23-JUL-1997;
                                                                                  23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-MAY-1999
              27-JAN-1999
                                                                                                                                                        Conrad B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Sim
Matches 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAW97745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAW97745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
```

```
This is the envelope protein (Env) of a new human endogenous retrovirus (HERV), designated IDDMK1.2-22, that has been identified as the source of superantigen (SAB) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAB activity within the env gene. A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV captivity (especially including the pre-symptomatic having SAB activity (especially IDDMK1.2-22 5'LTR, 3' long terminal repeat. env or pol); (b) a protein or peptide expressed HERV see AMB/7745-446); (c) antibodies specifically associated with the HERV. Products of the invention can be used to identify substances capable of blocking transcription or translation of SAP encoding nucleic acid sequences, useful in therapy and/or prevention of autoimmune disease associated with the SAB. A nucleic acid encoding the endogenous SAB in IDDM suggests a general model according to which self SAB-driven and systemic activation of autoreactive T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New isolated human endogenous retrovirus · used to develop product. for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27.9%; Score 840; DB 20;
100.0%; Pred. No. 1.4e-73;
iive 0; Mismatches 0;
                                                  17..19
/note= "Asn is N-glycosylated"
                                                                                                                           "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
Location/Qualifiers
17..19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 6; Fig 7D; 165pp; English.
                                                                                                                                                                                                                                                                                                                                              97EP-0401773
97EP-0112482
                                                                                                                                                                                                                                                                                           98WO-EP04926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAX07188, AAX07189.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                               42..45
/note=
                                                                                                                                                                                                                                                                                                                                                                                                                               (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-143118/12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         153 AA;
                                            Modified-site
                                                                                                     Modified-site
                                                                                                                                                                                   WO9905527-A2
                                                                                                                                                                                                                                                                                           22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                 23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                             22-JUL-1997;
                                                                                                                                                                                                                                      04-FEB-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Sin
Matches 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conrad B,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
```

ö

a

a

ô

g ò

õ

us-09-490-700-38.rag

ø

```
The sequence is that of all insulin depondent disperse mentious secondated human endogenous retrovirus (IDDMK1.2-22) env/fs protein. The retrovirus has Superantigen (SAg) activity. It can be used as part of a method is specifically used to diagnose type I diabetes mellitus. Modified proteins expressed by the retroviral sequence (without SAg activity but still able to induce an immune response) are useful in vaccines to treat or prevent SAg-related autoimmune classas; nucleic acid sequences encoding (modified) SAg can be used similarly to treat such diseases. Retroviral-encoded SAg are important in pathogenesis of autoimmune disease, probably by activating cutoreactive T cells. The method is very specific (it can differentiate between expressed and non-expressed viral nucleic acids) and can be used even where the pathogen is an ubiquitous endogenous retrovirus. Blood corp plasma samples can be tested without extensive preparation and diagnosis can be made before clinical signs are apparent, allowing carly intervention before severe tissue damage has occurred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Diagnosing human autoimmune disease by detecting retrovirus with superantigen activity - new retrovirus associated with type 1 diabetes, its proviral DNA, and related vectors, transformed cells, proteins, antibodies and specific binding agents, used for treating or preventing autoimmune disease
                                                                                                                                                                                            IDDM; insulin-dependent diabetes mellítus; endogenous retrovirus;
SAg; superantigen; provirus; autoimmune disease; type 1 diabetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sequence is that of an insulin-dependent diabetes mellitus
                                                                                                                                                      Human endogenous retrovirus IDDMK1.2-22 env/fs (sag) protein.
                                    AAW95694 standard; Protein; 181 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 31; Fig 7G; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                             97EP-0401773
                                                                                                                                                                                                                                                                                                                                                                                                                                  97EP-0401773
                                                                                                               (first entry)
                                                                                                                                                                                                                                       diagnosis; env; envelope.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-097928/09.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MACH/) MACH B F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAX07516.
                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                             23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                  23-JUL-1997;
                                                                                                                 08-JUN-1999
                                                                                                                                                                                                                                                                                                              EP893691-A1
                                                                                                                                                                                                                                                                                                                                                       27-JAN-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                             AAW95694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The
                    AAW95694
```

```
ó
                                                                                                                              61 LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                                                                                                             Gaps
                                                              1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC 60
                                 .
0
27.9%; Score 840; DB 20; Length 181;
100.0%; Pred. No. 1.8e-73;
ive 0; Mismatches 0; Indels (
                                                                                                                                                                                           PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                Matches 153; Conservative
                Similarity
                    Local
                                                                                                                                                           61
                                                                                                                                                                                           121
```

g à q ò

181 AA;

Seguence

```
121 pkgskntevlvweecvansvvilgnnefgtiid 153
  g
```

AAW97747 standard; Protein; 181

AAW97747 RESULT

Human endogenous retrovirus IDDKK1.2-22 Env/F-S (SAg).

(first entry)

21-MAY-1999

AAW97747;

```
This is the envelope (Env)/FS polypeptide of a new human endogenous retrovirus (HERV), designated IDDMK1.2-22, that has been identified as the source of Superantigen (SAG) activity in insulin-dependent diabetes mellitus (IDDM) patients. The endogenous retrovirus is ubiquitous in the human genome but is only expressed in diabetic individuals. The HERV encodes SAG activity within the env gene.

A claimed process for the diagnosis, including the pre-symptomatic diagnosis, of a human autoimmune disease associated with a HERV capagnosis, of a human autoimmune disease associated with a HERV having SAG activity (especially IDDMK1.2-22 5'LTR, 3' long terminal biological sample either: (a) the mRNA of an expressed HEV having SAG activity (especially IDDMK1.2-22 5'LTR, 3' long terminal repeat, env or pol); (b) a protein or peptide expressed by the HERV (see AAW97745-48); (c) antibodies specific to the proteins expressed by the HERV. Products of the invention can be used to identify substances capable of blocking transcription or translation of SAG-encoding nucleic acid sequences, useful in therapy and/or prevention of autoimmune disease associated with the SAG. A nucleic acid encoding the endogenous SAG in IDDM suggests a general model according to which self SAG-driven and systemic activation of autoreactive T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New isolated human endogenous retrovirus - used to develop products for the diagnosis, prevention and treatment of autoimmune disease, particularly insulin dependent diabetes mellitus
                                                                                                         HERV; IDDKKI.2-22; superantigen; SAg; antigen; IDDM; insulin-dependent diabetes mellitus; autoimmune disease; diagnosis; therapy; vaccine; envelope protein; env gene.
                                                                                                                                                                                                                                       17..19
/note= "Asn is N-glycosylated"
                                                                                                                                                                                                                                                                                        /note= "Asn is N-glycosylated"
                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 6; Fig 7G; 165pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                   98WO-EP04926.
                                                                                                                                                                                                                                                                                                                                                                                                                                   97EP-0401773.
97EP-0112482.
                                                                                                                                                                                  Human endogenous retrovirus
                                                                                                                                                                                                                                                                         42..45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-143118/12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MEDI-) MEDIGEN SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conrad B, Mach B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAX07191
                                                                                                                                                                                                                                                                       Modified-site
                                                                                                                                                                                                                                       Modified-site
                                                                                                                                                                                                                                                                                                                             WO9905527-A2
                                                                                                                                                                                                                                                                                                                                                                                                   22-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                     23-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      22-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                               04-FEB-1999
```

```
Matches
                                                                                                                                                                                                                                                                                                                                                      AAY41142
  8×88888888888
                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                             ò
                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention describes primer sets for synthesising 5602 full-length CDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the 5'-end sequence'3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynuclectides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the
                                                                                                              LMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKTCPKEI 120
                                                 9
                                                                                                                                                                                                                                                                                                                                       Human; primer; detection; diagnosis; antisense therapy; gene therapy.
                           Gaps
                                                             Yamamoto J;
                                                1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGC
                          ö
Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Saito K, Ya
 27.9%; Score 840; DB 20; 100.0%; Pred. No. 1.8e-73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 8; SEQ ID 12136; 2537pp + CD ROM; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hayashi K, S
A, Nagai K,
       100.0%; Preu. ....
                                                                                                                                                121 PKGSKNTEVLVWEECVANSVVILQNNEFGTIID 153
                                                                                                                                                              121 pkgskntevlvweecvansvvilgnnefgtiid 153
                                                                                                                                                                                                                                                                                                                 Human protein sequence SEQ ID NO:12136.
                                                                                                                                                                                                                                       AAB93186 standard; Protein; 141 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sugiyama T, Wakamatsu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nishikawa T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                99JP-0248036.
99JP-0300253.
2000JP-0118776.
2000JP-0183767.
2000JP-0241899.
                                                                                                                                                                                                                                                                                                                                                                                                                                         28-JUL-2000; 2000EP-0116126
                                                                                                                                                                                                                                                                                        (first entry)
                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (HELI-) HELIX RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-318749/34.
           Local Similarity
nes 153; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Isogai T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-JAN-2000;
02-MAY-2000;
09-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                         EP1074617-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                29-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                07-FEB-2001
                                                                                                                                                                                                                                                                                       26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ś
                                                                                                                                                                                                                                                                AAB93186
 Query Match
              Best Loca
Matches
                                                                                                 61
                                                                                                 ö
                                                                                                                        음
                                                                                                                                                  ò
                                                                                                                                                                         g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention provides a method for inducing an immune response to mouse mammary tumor virus (MMTV) in a human subject. The method comprises administering to the subject an immunomodulatory composition comprising a
the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH33166 to AAH13628 and AAH13631 to AAH13642 represent human cDNA sequences; AAB92466 to AAH3639 represent human amino acid sequences; and AAH13629 to AAH33632 represent oligonucleotides, all of which are used in the exemplification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pharmaceutical carrier and at least one MMTV antigen (or a discontinuous epitope of MMTV) in an amount sufficient to elicit an increase in a MMTV specific cellular or humoral response. The compositions and methods of the invention are useful for treating or preventing a human disease or discorder caused by MMTV, or a biological or chemical agent that reacts with an antibody directed towards an MMTV antigen. The compositions and antibodies may be used in passive immunotherapy. The immunomodulatory
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               compositions can be used in a preventative manner for those subjects not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Novel immunomodulatory compositions of mouse mammary tumor virus (MMTV) antigens used for treating or preventing MMTV infections
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  299 LFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIFT 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 359 LIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQINDL 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          response; mouse mammary tumor virus; MMTV; immunomodulatory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                      Length 141;
                                                                                                                                                                                                                                                                                                               Score 557; DB 22; Length 1
Pred. No. 4.2e-46;
4; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mammary tumor virus (MMTV) env polypeptide
                                                                                                                                                                                                                                                                                                                                                                                     14; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              epitope; passive immunotherapy; env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY41142 standard; Protein; 688 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (ITBI-) INT BIOIMMUNE SYSTEMS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Fig 5; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        419 ROTVIWMGDRLDLEHHFQL 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tch 18.5%; al Similarity 77.0%; 107; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-US07712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0057003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mouse mammary tumor virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           بر
ڊ,
                                                                                                                                                                 of the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gershoni
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-620171/53.
                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                             141 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAZ23193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09951268-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Stewart THM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14-0CT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY41142;
                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Immune
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mouse
```

æ

```
Sequence
                                                                    Fan HY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 13
                                                                                                                                                                                                                                                                                                                 55
                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                                                                                                                                                                                                 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      298
                                                                                                                                                                                                                                                                                                                                                                                                             222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR31349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ð
   q
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                               Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                     25;
exposed to MMTV or MMTV antigen reactive agents. The MMTV antigens also have use in diagnostic assays. The present sequence represents the amino acid sequence of the MMTV env polypeptide.
                                                                                                                                                                                 164
                                                                                                                                                                                               193 g------kgdkrrmwelwlttlgnsga---ntklvpikklp--pkyphcgiafkkda 239
                                                                                                                                                                                                                      185
                                                                                                                                                                                                                                                                                                           ------krpgfqehem----iptsacvtypyaillglpglidiekrgstfhiscssc 401
                                                                                                                   --GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPK 112
                                                                                                                                                                                                                                      240 fwegdesapprwlpcafpdggvsfspkgalgllwdfslpspsvdgsdgiksknlfgnyt 299
                                                                                                                                                                                                                                                                     ::| |: :: :: | || | : : Dkdpndrdftalvphtelfrlvaasrhlil-- 354
                                                                                                                                                                                                                                                                                              297
                                                                                                                                                                                                                                                                                                                                   RLFTCIDSTFNWQHRILLVRAREGMWIPVST-DRPW---EASPSIHILTEILKGVLNRSK 353
                                                                                                                                                                                                                                                                                                                                                                       RFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLAS 413
                                                                                                                                                                                                                                                                                                                                                                                                            QINDLRQTVIWMG-DRLDLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQG--R 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                        4 PVTW-MDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGYHYP-----PICL-- 54
                                                                                                                                                                                GKTCPKEIPKGSKNTEVLVWEECVANSVVILQNNEFGTIIDXAPRGQFYHNC------
                                                                                                                                                                                                                                                                                                                                              ---SDLT
                                                                                                                                                                                                                                                                                              GNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYMLVVG-----NIVIKPASQTITCENC
                                                                                                                                                                                                                                                                                                                                                                                  EDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTIRSTMIINLI
                                                                                                                                                                                                                                                         ESLDKHKHKKLQSFYL - - - WEWEEKGISTPRPKIISPVSGPEHPELWRLTVASHHIRIWS
                                                                                     152;
                                                                  Length 688;
                                                                                     217; Indels
                                                                                                                                                                                                                     ---AQVSPAVD----
                                                                  Score 398; DB 20;
Pred. No. 1.5e-29;
                                                                                     Mismatches
                                                                                                                                                            sfstptgcfgvdkgvflsdtptvdnn------
                                                                  13.2%; Score 398; 23.7%; Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      615
                                                                                     100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cytostatic; lung cancer; ENV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB66969 standard; protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jaagsiekte sheep retrovirus
                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LIVVCLFCLLLVCRC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |:|: | :| :|
|lvivlmifpivfqc 648
                                                                                                                                                                                                                    165 --SGQTQSCPS----
                                                                            Similarity
                                     688 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JSRV ENV protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200104266-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-APR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-JAN-2001.
                                                                            Best Local Sim:
Matches 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB66969;
                                       Sequence
                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        634
                                                                                                                                                                                                                                                                                              243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      531
                                                                                                                                            55
                                                                                                                                                              164
                                                                                                                                                                                113
                                                                                                                                                                                                                                                         186
                                                                                                                                                                                                                                                                           300
                                                                                                                                                                                                                                                                                                                 355
                                                                                                                                                                                                                                                                                                                                   298
                                                                                                                                                                                                                                                                                                                                                     402
                                                                                                                                                                                                                                                                                                                                                                       354
                                                                                                                                                                                                                                                                                                                                                                                         456
                                                                                                                                                                                                                                                                                                                                                                                                            414
                                                                                                                                                                                                                                                                                                                                                                                                                                                471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 1
                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2222×8
                                                                                                                                           ò
                                                                                                                                                            g
                                                                                                                                                                                ò
                                                                                                                                                                                              Q
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
The present invention relates to a replication competent infectious dagagiskite sheep retrovirus (JSRV). The present sequence is the env protein of the JSRV of the present invention. The JSRV of the present invention may be used to treat a cell proliferative disorder in a mammal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                115 TCPKEIPKGSKNTEVLV----WEECVANSVVILQ-NNEFGTIIDXAPRGQFYHNCSGQT 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----flwdwapgg--tpdfpeyrgqhppifsvn 248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  392 YWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITP 450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |: : | : |||||: : | :|
92 miqslgwdreivpvyvndtsllggksd--ihispqq----anisfyglttgy-pmcfsyg 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGK 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                145 sqhphciqvsad---isyprvt----isqid-----ektgkksygng 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 tapiyqtelwkllaafgh----gn----slylqpnisgskygdvgvtgfl-yp-racvp 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         274 PPYMLVVGNIVIKPASQT - - ITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRP 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            332 WEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVN 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         416 slsynvtkvmgtgedidkkiedrlsalydvvrvlgegvgsinfrmkigchanykwicvtk 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINIS---IGYHYPPICL---54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : | :|| |: : : | | | : : | | sgpldipfcdkhlsigigidtpwilcrarvasvyninnanat---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                169 QSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTP-----RPKIISPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                221 SGPEH-PELWRLTVASHHIRIWSGNQTLETRYRKP-----FYTIDLNSILTVPLQSCLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ypfmligghmeitlslniyhlncsnciltncirgvakge-gviivkgpafvmlpveiaea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QIYNESEHHWDMVRRHLQG--REDNLTLDISKLKEQIFE---ASKAHLNLVPGTE-AIAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 134;
                                                                                                                                                                                                                                                                                                                                                                   Isolated replication competent infectious Jaagsiekte sheep ret.
(JSRV) used to treat a cell proliferative disorder in a mammal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 11.5%; Score 345; DB 22; Best Local Similarity 22.9%; Pred. No. 1.8e-24; Matches 132; Conservative 108; Mismatches 203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 505 VADGLANLNPVTWIKTIRSTMIINLILIVVCLF-CLL 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR31349 standard; Protein; 615 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 3; Fig 6; 123pp; English.
                                                                                                                                                                                                                       Ä
38-JUL-2000; 2000WO-US18856
                                                                       99US-0142868
                                                                                                                                               (REGC ) UNIV CALIFORNIA
                                                                                                                                                                                                                       Palmarini M,
                                                                                                                                                                                                                                                                                               WPI; 2001-112617/12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      615 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  e.g. lung cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------
                                                                    38-JUL-1999;
```

28;

σ

```
The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the
                                                                                                                                                                                                                                                                                                                                                                  | |:: :| ||:| |:: | 4:| | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 
Human; primer; detection; diagnosis; antisense therapy; gene therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ņ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            298 ypfmligghmeitlslniyhlncsnciltncirgvakge-gviivkgpafvmlpvelaea
                                                                                                                                                                                                        332 WEASPSIHILTEILKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVN
                                                                                                                                                                                                                                                                    wydetalellgri-ntalsrpkrglsliilgivslitliatavtacvslagsigaahtvd
                                                                                                                                                                                                                                                                                                                                392 YWQKNSTRLWNSQSSIDQKLASQINDLRQTVIWMGDRL-DLEHHFQLQCDWNTSDFCITP
                                                                                                                                                                                                                                                                                                                                                                                                                                                    451 QIYNESEHHWDMVRRHLQG--REDNLTLDISKLKEQIFE---ASKAHLNLVPGTEAIAGV
                                                                                       PPYMLVVGNIVIKPASQT - - ITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yamamoto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Saito K,
Otsuki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             506 ADG-LANL-NPVTWIKTIRSTMI---INLILIVVCLFCLLLVCR 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 8; SEQ ID 18882; 2537pp + CD ROM; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hayashi K, S
A, Nagai K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human protein sequence SEQ ID NO:18882.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nishikawa T,
T, Wakamatsu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB95842 standard; Protein; 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29-JUL-1999; 99JP-0248036.
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
02-MAY-2000; 2000JP-0183767.
09-JUN-2000; 2000JP-0241899.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-2000; 2000EP-0116126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (HELI-) HELIX RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-318749/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sugiyama
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Isogai T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP1074617-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-FEB-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ota T, I
Ishii S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB95842;
                                                                                                                                                                                                                                                                       357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                       δλ
                                                                                                                                             g
                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sghphciqvsad---isyprvt----isqid-----ektqkksyqnq 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    115 TCPKEIPKGSKNTEVLV----WEECVANSVVILQ-NNEFGTIIDXAPRGQFYHNCSGQT 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            221 SGPEH-PELWRLTVASHHIRIWSGNQTLETRYRKP-----FYTIDLNSILTVPLQSCLK 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55 GRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGK 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JSRV causes epithelial carcinoma in ovine animals, partic. pulmonary adenomatosa in sheep. The complete cDNA sequence of the JSRV genome was prepared from an approx. B.7kb band of poly-A RNA isolated from seemi-purified lung lavage samples from infected sheep. The invention includes the Env amino acid sequence or any part of it which is capable of specific immunological reaction with antibodies directed against JSRV. The glycoproteins gp46 or gp31 and the precursor PR69 env are preferred. See also AAR31346-R31348 and AAQ35153-Q35155.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTP-----RPKIISPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINIS---IGYHYPPICL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tgpldipfcdkhlsigigidtpwtlcrarvasvyninnanat-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New Jaagslekte Retrovirus and characteristic nucleic acid · also derived proteins, probes and antibodies, useful for in vitro diagnosis and in vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 615;
                                                                                                                                                                                                        epithelial carcinoma; ovine; sheep; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 340.5; DB 14;
Pred. No. 4.9e-24;
); Mismatches 202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "transmembrane portion"
                                                                                                                                                                                                                                        pulmonary adenomatosa; envelope glycoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      York D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (INRM ) INSERM INST NAT SANTE & RECH MED.
                                                                                                                                                                                                                                                                                                                                                                                                     /note= "surface portion" 379..615
                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 26; Page 41-43; 75pp; French.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vigne R,
                                                                                                                                                Jaagsiekte retrovirus Env protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11.3%;
23.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      91FR-0006060.
                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Verwoerd D,
                                                                                                                                                                                                                                                                                                Jaagsiekte retrovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1993-020250/03.
N-PSDB; AAQ35153.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          615 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-MAY-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-MAY-1991;
                                                                                 18-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FR2676455-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20-NOV-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Querat GF,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 139;
                            AAR31349;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                            Key
Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180
```

a

ò

QQ

ò

a õ

ò

ò

```
ö
polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs as as a second of the full-length AAH13632 to AAH13638 and AAH13633 to AAHH3642 represent human cDNA sequences; ABB92446 to AAH13632 represent human amino acid sequences; and AAH13632 to AAH13632 represent oligonucleotides, all of which are used in the exemplification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          462 MVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWIKTI 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8.6%; Score 260; DB 22; Length 106; 59.0%; Pred. No. 2.1e-17; Live 17; Mismatches 17; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     522 RSTMIINLILIVVCLFCLLLVCR 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 59.09
Matches 49; Conservative
                                                                                                                                                                                                                                                                                          of the present invention
                                                                                                                                                                                                                                                                                                                                         106 AA;
                                                                                                                                                                                                                                                                                                                                           Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs. The primers allow obtaining of the full-length AAH13613 to AAH13612 represent human anino acid sequences; AAH36446 to AAH13612 represent human anino acid sequences; and AAH13629 to AAH13612 represent coligonucleotides, all of which are used in the exemplification

the present invention.

276 AA;

Seguence

the 5602 nucleotide sequences defined in the specification, where the obliganciectide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end

complementary strand of a polynucleotide which comprises one of

8

Gaps

```
||: : | | :: : |:||||: || ||: ::| : :| : | hcinhstlgthnistlmilgcipglwipvnlsepwaatialhfvkllltgfthcvrrglg ,215
                                                                                            149 GTIIDXAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKKKLQSFYLWE---- 203
                                                                                                                                                                                        ---WEEKGISTPRPK-IISPVSGPEHPELWRLTVASHHIRIW---- 241
                                                                                                                                                                                                                                                                                 SGNQTLETRYRKPFYTIDLNSILTVPLQSCLKPPYM-LVVGNIVIKPASQTITCENCRLF 300
                                                                                                                                                                                                                                                                                                                        100 snnsnskqhslnvtfvknittqftv----cvfnpyvflaakkdqlqvnntqltckscqly 155
                                                                                                                                                                                                                                                                                                                                                                             301 TCIDSTFNWQHRI----LLVRAREGMWIPVSTDRPWEASPSIHILTEILKGVLNRSKRFIF 357
                                                                                                                                                                                                             40 slyntgiqsqsaaqiawhgagfspplpqwhylgrkgplqkmiwkaafpfmngniwvaiil 99
                                                                                                                                     4 gdivdwgpkgql------dgkeengkswhk--lcwhwwqafnas 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNYWQKNSTRLWNSQSSIDQKLASQ1 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.4%; Score 222.5; DB 22; Length 276; 21.5%; Pred. No. 4.2e-13; Artive 58; Mismatches 117; Indels 59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Search completed: April 9, 2002, 16:56:44
                                                Conservative
Query Match
Best Local Similarity
Matches 64; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Job time: 478 sec
                                                                                                                                                                                                                                                                                 242
                                                                                                                                                                                     204
                                                                                                                                                                                                                                                                                                                                                                                                                       156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   216
                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                           οq
                                                                                                                                       g
                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -
                                                                                                                                                                                                                                                                                                                                                                                                 Human; primer; detection; diagnosis; antisense therapy; gene therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yamamoto J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Saito K,
, Otsuki'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hayashi K, S
A, Nagai K,
                                                                                                                                                                                                                                                                                                                                                   Human protein sequence SEQ ID NO:18817
                                                                                                                                                                                                             AAB95817 standard; Protein; 276 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ogai T, Nishikawa T,
Sugiyama T, Wakamatsu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-JUL-1999; 99JP-0248036.
27-ANG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
02-MAX-2000; 2000JP-0188776.
09-JUN-2000; 2000JP-0241899.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-2000; 2000EP-0116126.
                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (HELI-) HELIX RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-318749/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Isogai T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EP1074617-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07-FEB-2001.
                                                                                                                                                                                                                                                                                                     26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ishii S,
                                                                                                                                                                                                                                                         AAB95817;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ė
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ota
                                                                                                                                                                                     AAB95817
```

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary

Claim 8; SEQ ID 18817; 2537pp + CD ROM; English.